

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2001, 23:48:11 ; Search time 1789.67 seconds
(without alignments)
6231.369 Million cell updates/sec

Title: US-09-735-712-9
676
Sequence: 1 tcatctccttcaattatc.....tgyttaaataaaaaaa 676

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Hit number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_da:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sy:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_hggo_hum:*
31: em_hggo_inv:*
32: em_hggo_rod:*
33: em_hggo_hum:*
34: em_hggo_inv:*
35: em_hggo_rod:*
36: em_hggo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	676	100.0	676	6	AX179789	AX179789 Sequence
2	667.2	98.7	694	9	AF237907	AF237907 Homo sapi
3	666.6	98.6	697	9	AF231127	AF231127 Homo sapi
4	665.6	98.5	691	9	AB013103	AB013103 Homo sapi
5	603	89.2	603	6	AX179781	AX179781 Sequence
6	347.4	51.4	450	6	AX179787	AX179787 Sequence
7	341.8	50.6	417	6	AX179785	AX179785 Sequence
8	246	36.4	246	6	AX179783	AX179783 Sequence
9	185.4	27.4	138097	2	AC027787	AC027787 Homo sapi
10	185.4	27.4	138097	2	AC027787	AC027787 Homo sapi
11	185.4	27.4	147788	2	AP0015840	AP0015840 Homo sapi
12	185.4	27.4	161039	2	AP0013127	AP0013127 Homo sapi
13	185.4	27.4	166804	2	AP001034	AP001034 Homo sapi
14	185.4	27.4	166804	2	AC090401	AC090401 Homo sapi
15	157	23.2	161039	2	AP0013127	AP0013127 Homo sapi
16	148.6	22.0	166804	2	AP001034	AP001034 Homo sapi
17	88	13.0	1076	9	AF068288	AF068288 Homo sapi
18	86.4	12.8	720	9	AB022821	AB022821 Homo sapi
19	86.4	12.8	916	9	AB013102	AB013102 Homo sapi
20	86.4	12.8	1619	9	AF237912	AF237912 Homo sapi
21	86.4	12.8	1669	6	AR035695	AR035695 Sequence
22	86.4	12.8	1669	6	AR083580	AR083580 Sequence
23	76.2	11.3	2805	9	AF367473	AF367473 Homo sapi
24	63.2	9.3	1546	9	HUMIERB	L35848 Homo sapien
25	63.2	9.3	1661	6	A68627	A68627 Sequence 3
26	63.2	9.3	1661	6	AR082014	AR082014 Sequence
27	63.2	9.3	1661	6	I25809	I25809 Sequence 3
28	62.4	9.2	1222	4	SSR236932	AJ236932 Sus scrofa
29	61.6	9.1	1628	9	BC008487	BC008487 Homo sapi
30	58.6	8.7	708	6	AX101306	AX101306 Sequence
31	58.6	8.7	1010	10	MUSFCERB	J05019 Mouse mast
32	55.6	8.2	899	10	AF237910	AF237910 Mus muscu
33	55.4	8.2	1163	6	AF237915	AF237915 Mus muscu
34	55.4	8.1	1474	6	I07272	I07272 Sequence 19
35	55	8.1	1474	9	HSCD20	X12530 Human mRNA
36	55	8.1	1476	6	I07277	I07277 Sequence 24
37	55	8.1	1482	9	BC002807	BC002807 Homo sapi
38	55	8.1	1597	9	HSCD20A	X07203 Human mRNA
39	54.8	8.1	518	10	AF280401	AF280401 Mus muscu
40	53	7.8	1480	10	AF237914	AF237914 Mus muscu
41	52.6	7.8	862	9	AB013104	AB013104 Homo sapi
42	52.6	7.8	930	9	AF286866	AF286866 Homo sapi
43	52.6	7.8	1131	9	AF212240	AF212240 Homo sapi
44	52.6	7.8	1152	9	AF237908	AF237908 Homo sapi
45	52.6	7.8	1289	9	AF142409	AF142409 Homo sapi

ALIGNMENTS-

RESULT 1
AX179789
LOCUS AX179789 676 bp DNA
DEFINITION Sequence 9 from Patent WO0146417.
ACCESSION AX179789
VERSION AX179789.1 GI:15132148
KEYWORDS
SOURCE
ORGANISM human.
Human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 676)
Wolke, D.W. and Turner, C.A.
TITLE Human membrane proteins and polynucleotides encoding the same
having homology to cd20 proteins and ige receptors
JOURNAL Patent: WO 0146417-A 9 28-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1..676
/organism="Homo sapiens"

BASE COUNT 193 a 140 c 112 g 231 t
ORIGIN

Query Match 100.0%; Score 676; DB 6; Length 676;
Best Local Similarity 100.0%; Pred. No. 9,5e-148;
Matches 676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 tcatctcttccaatttcacccgacacatcatgatgattcaagcagccagacagtcg 60
DB 1 TCATCTCTCTTCAAAATTTTCACCGACACCATCATGATTCAGCACCGCACATCCGG 60
QY 61 ttttcttgatcttctccacgaatcactgcttcagaaatgagtcacagaacttcag 120
DB 61 TGTTCCTGATTTCTCTCAGAAATCACTGCTTCAGAAATGAGTCACAGAACTTCAG 120
QY 121 ccaagaccttccaactaaagccctgtgcaaaaattttgtctagaaaatgaaatct 180
DB 121 CCAGCACTTTTCAACTAAAGCCCTTGCAAAAATTTTGTCTAGAAAATGAAAATCT 180
QY 181 taggactacacagatccctgttgtaattagaccttctcttgagatctctcctt 240
DB 181 TAGGACTATCCAGATCCGTTTGGAATTTAGACTTTCTTTTGAGATTATCTCCCTT 240
QY 241 tcacctgttaaacacatataccaaagttcccttatacttcttcaggataccatct 300
DB 241 TCACCTGTGTAAACCAATCCAAAGTTCCCTTATTTCTTTTCAGGATTCATCTCT 300
QY 301 ggggctcgtttgttcttaattcttgagaccttccaaatgcaatgaaagaaacca 360
DB 301 GGGGCTCTGTTTGTCTTAATTTCTGAGCCTTCTTAATGCACTGAAAGAAACCA 360
QY 361 cagaacctcgaataatctgagccgaataatgaatctctctagtcgcttgagaagatag 420
DB 361 CAGAAACTGTATATATTTGAGCCGAATGAAATCTTCTAGTCCCTTGAGAGCAATAG 420
QY 421 ctggaatcatctctcctcactttgtttcaatctcctagatcaaaaactaatgtgtatt 480
DB 421 CTGGAATCATCTCTCCTCAATTTGGTTTCATCTAGATCAAAACATCTTGTGTTATT 480
QY 481 ctcccaaaatagtcagtgtaagcgtgttactgtctgtctcttgagaaatttgatcat 540
DB 481 CTCACCAAAATAGTCAGTAAAGGCTGTACTGTCTGTGGGAAATTTGATTAAT 540
QY 541 tgaatcattcaagaatctgaatcattcattctcctccttcaatttgagggtgc 600
DB 541 TGATGACTTTCAGCATTTATGATTAATTCATTTCTGCTTTCATTTTGGGTCCT 600
QY 601 actcagagatctgtgattgtgaaacattgttgactagcactgtgagaataaagatgtg 660
DB 601 ACTCAGAGATTTGTGATTTGTAACAATGTGTGTGACTGACACTGAGAAATTAAGATGTG 660
QY 661 ttaataataaaaaaa 676
DB 661 TTAATAATAAAAAAA 676

RESULT 2
LOCUS AF237907 694 bp mRNA PRI 17-Apr-2001
DEFINITION Homo sapiens MS4A5 protein mRNA, complete cds.
ACCESSION AF237907
VERSION AF237907.1 GI:13649400
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 694)
AUTHORS Liang,Y. and Tedder,T.F.
TITLE Identification of a cd20-, fcepsilonibeta-, and htm4-related gene
family: sixteen new ms4a family members expressed in human and

```

Mouse
JOURNAL Genomics 72 (2), 119-127 (2001)
MEDLINE 21295030
REFERENCE 2 (bases 1 to 694)
AUTHORS Liang,Y. and Tedder,T.F.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-2000) Immunology, Duke Medical Center, Research
Dr., Durham, NC 27710, USA

FEATURES
source location/Qualifiers
1..694
/organism="Homo sapiens"
/db_xref="taxon:9606"
55..657
/product="MS4A5 protein"
/protein_id="AAK37416.1"
/db_xref="GI:13649401"

CDS

/translation="MDSSTAHSPVFLVFPPEPTASVESEFELSATTSSTOSPLQKLF
RKMILGTIQLFIMTFSEGVILFLLKYPFPPTFLSGPFWGVLEFINSGL
IAVRKRTETLILSRIMNLSALGAILIFGLIDONYICGYSHONSCKAVT
VLFGLITMTFSIIELFISLPSIIGSEDDCEOC"

BASE COUNT 198 a 143 c 116 g 237 t
ORIGIN

Query Match 98.7%; Score 667.2; DB 9; Length 694;
Best Local Similarity 99.6%; Pred. No. 1.1e-145;
Matches 669; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 tcatctcttccaatttcacccgacacatcatgatgattcaagcagccagacagtcg 60
DB 23 TCATCTCTCTTCAAAATTTTCACCGACACCATCATGATTCAGCACCGCACATCCGG 82
QY 61 ttttcttgatcttctccacgaatcactgcttcagaaatgagtcacagaacttcag 120
DB 83 TGTTCCTGATTTCTCTCAGAAATCACTGCTTCAGAAATGAGTCACAGAACTTCAG 142
QY 121 ccaagaccttccaactaaagccctgtgcaaaaattttgtctagaaaatgaaatct 180
DB 143 CCAGCACTTTTCAACTAAAGCCCTTGCAAAAATTTTGTCTAGAAAATGAAAATCT 202
QY 181 taggactacacagatccctgttgtaattagaccttctcttgagatctctcctt 240
DB 203 TAGGACTATCCAGATCCGTTTGGAATTTAGACTTTCTTTTGGAATTTATCTCTCT 262
QY 241 tcacctgttaaaaacacatataccaaagttcccttatacttcttcaggataccatct 300
DB 263 TCACCTGTGTAAACCAATCCAAAGTTCCCTTATTTCTTTTCAGATTAATCATCTCT 322
QY 301 ggggctcgtttgttcttaattcttgagaccttccaaatgcaatgaaagaaacca 360
DB 323 GGGGCTCTGTTTGTCTTAATTTCTGAGCCTTCTTAATGCAATGAAAGAAACCA 382
QY 361 cagaacctcgaataatctgagccgaataatgaatctctctagtcgcttgagaagatag 420
DB 383 CAGAAACTGTATATATTTGAGCCGAATGAAATTTCTTAGTCCCTTGAGAGCAATAG 442
QY 421 ctggaatcatctcctcacttggtttcaatcctcctcctccttcaatttgagggtgc 480
DB 443 CTGGAATCATCTCTCCTCAATTTGTTTCATCTGATCAAAACATTTGTGTTATT 502
QY 481 ctcccaaaatagtcagtgtaagcgtgttactgtctgttcttggaatttgatcat 540
DB 503 CTCACCAAAATAGTCAGTAAAGGCTGTACTGTCTGTGGGAAATTTGATTAAT 562
QY 541 tgaatcattcaagaatctgaatcattcattctcctccttccaaatttgagggtgc 600
DB 563 TGATGACTTTCAGCATTTATGATTAATTCATTTCTGCTTTCATTTTGGGTCCT 622
QY 601 actcagagatctgtgattgtgaaacattgttgactagcactgtgagaataaagatgtg 660
DB 623 ACTCAGAGATTTGTGATTTGTAACAATGTGTGTGACTGACACTGAGAAATTAAGATGTG 682

```

```

QY 661 ttaataaaaaa 672
Db 683 TTTAAATVAAAAA 694

RESULT 3
AF321127 697 bp mRNA PRI 07-FEB-2001
LOCUS Homo sapiens testis-expressed transmembrane-4 protein (TMEM4) mRNA,
DEFINITION complete cds.
ACCESSION AF321127
VERSION AF321127.1 GI:12698681
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 697)
Hulet,M.D., Pagler,E., Hornby,J.R., Hogarth,P.M., Eyre,H.J.,
Baker,E., Crawford,J., Sutherland,G.R., Ohms,S.J. and Parish,C.R.
Isolation, tissue distribution, and chromosomal localization of a
novel testis-specific human four-transmembrane gene related to CD20
and FcεpsilonRI-beta
Biochem. Biophys. Res. Commun. 280 (1), 374-379 (2001)
2 (bases 1 to 697)
Hulet,M.D.
Direct Submission
Submitted (14-NOV-2000) Division of Immunology and Cell Biology,
John Curtin School of Medical Research, Mills Road, Canberra, ACT
2601, Australia
FEATURES
source 1..697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q12"
1..697
/gene="TMEM4"
57..659
/gene="TMEM4"
/codon_start=1
/product="testis-expressed transmembrane-4 protein"
/db_xref="GI:12698682"
/translation="MDSSTANSPVPLVPPPTTASEYSTEISTESTOSTSLQKLF
RKMKLGIQLFGIMTSEFVLEFLTKPRPRPFILSGYPPWGSVLFINSGLF
IAVKRTETELIISRIMNPLSALGAIIGILTPGFLIDONYICGYSHONSQCKAVT
VLLGLITLMTFSLIEFLISLPSILGCHSDCDEQCC"
CDS
gene
CDS
BASE COUNT 199 a 142 c 119 g 237 t
ORIGIN
Query Match 98.6%; Score 666.6; DB 9; Length 697;
Best Local Similarity 99.4%; Pred. No. 1.5e-145;
Matches 669; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 tatctcccttcaaatatcaccgacacatcatggtatccaagcaccgacacagtcgg 60
Db 25 TCATCCCTTCAAAATATACACGACCATCATGATTCAGCACACGACACAGTCGCG 84
QY 61 tcttctggtattctcctcagaataactgctcagatgatgagtcacagaaacttcag 120
Db 85 TGTTCCTGTATTTCTCCAGAAATCAGTCCTTCAGAAATATGATGCCAGAACTTCAG 144
QY 121 ccagcagcttccaactcaagcccttcgcaaaatatttgcctagaaaaatgaaatc 180
Db 145 CCACGACCTTTTCAACTCAAAAGCCCTTGCAAAAATTATTGCTAGAAAAATGAAAATCT 204
QY 181 taggagactacagatccggttggaatatgacctttcttggagtgatctccctt 240
Db 205 TAGGGACTATCCAGATCCGTTTGGAATATGACCTTTCTTTGGAGTTATCTCCCTT 264

```

```

QY 241 tcaactgttaaaaccatccaaggttcccttattatcttcttcaagatcattct 300
Db 265 TCACCTTTGTTAAACCATATCCAAAGCTTCCCTTATATTCTTTTCAGGATATCCATTCT 324
QY 301 ggggctctgtttgttccatcaatctcggagccttcctaattgagtgagtgaaaaaacc 360
Db 325 GGGGCTCTGTTTGTTCATTAATTCGAGCCCTTCCATTAATGCAAGTCAAAAAAGCA 384
QY 361 cagaactctgataatatgagccggaataatgatatcttctttagtcctcgaagcaatag 420
Db 385 CAGAAACCTCGATTAATATGAGCCGGAATTAATGAAATTTCTTAATGCCCCTGGAGCAATAG 444
QY 421 ctggaatcattccctccacattggtttcctcctagatcaaaactacattggtatt 480
Db 445 CTGGAATCATTTCTCCTCACATTTGGTTTCATCTAGATCAAACTACATTGTGGTTATT 504
QY 481 ctaccaaaatagtcagtgtaagcctgttactgttcctgtctcttggaattgattacat 540
Db 505 CTCACCAAAATAGTCAGTGAAGCGTGTACTGCTGCTGTTGGGAATTTGATTACAT 564
QY 541 tgaagtctcagcatattgaaatattatctctctgtccttcctaatttgggtgc 600
Db 565 TGATGACTTTCACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 624
QY 601 actcagagatttgatttggaacaatggttgactagacactgtagaataagaatgtg 660
Db 625 ACTCAGAGATTGTGATGATGGAACAATGTTGTGACTAGCAGCTGTGGAATTAAGATGTG 684
QY 661 ttaataaaaaa 673
Db 685 TTTAAATVAAAAA 697

```

```

RESULT 4
AB013103
LOCUS AB013103 691 bp mRNA PRI 20-MAR-2001
DEFINITION Homo sapiens mRNA for MS4A5, complete cds.
ACCESSION AB013103
VERSION AB013103.1 GI:11559213
KEYWORDS MS4A5; CD20-like 2.
SOURCE Homo sapiens testis cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (sites)
Ishibashi,K., Suzuki,M., Sasaki,S. and Imai,M.
Identification of a new multigene four-transmembrane family (MS4A)
related to CD20, HTM4 and beta subunit of the high-affinity IGE
receptor
JOURNAL Gene 264 (1), 87-93 (2001)
MEDLINE 21142397
REFERENCE
2 (sites)
Ishibashi,K., Sasaki,S. and Marumo,F.
Cloning of three CD20 homolog from human, putative calcium channels
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 691)
AUTHORS Ishibashi,K.
TITLE Direct Submission
JOURNAL Submitted (20-APR-1998) Kenichi Ishibashi, Tokyo Medical and Dental
University, 2nd Internal Medicine, Yushima 1-5-45, Bunkyo, Tokyo
113-8519, Japan (E-mail:Kishibashi.med2med.tmd.ac.jp,
Tel:81-3-5803-5223, Fax:81-3-5803-0132)
FEATURES
source 1..691
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="testis"
52..654
/gene="MS4A5"
52..654
/gene="MS4A5"
/codon_start=1
/product="MS4A5"

```

/protein_id="BAB18739.1"
/db_xref="GI:1153214"
/translation="MDSSTAHPEVLEPPPEITASEYESTELSATSTESQPLQKFLA
RKMILGILQILFIMFPEFVLEFLKRPPEFLEFGEYFPMGVLEFINSQAF
IAVRKTEETLIIISRIINMELSAIGAILLEFGLFDLQNYICGYSHNSQKAVT
VLFGILITLMTFSEIIEFLPSILGCHSEDDCECC"

BASE COUNT 195 a 142 c 117 g 237 t
ORIGIN

Query Match 98.5%; Score 665.6; DB 9; Length 691;
Best Local Similarity 99.4%; Pred. No. 2.5e-145;
Matches 668; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 tcaatccttcaaatatcaccgacacacatcatgatcaagacacgacacagtcg 60
DB 20 TCATCTCTTCAAAATATACCGACACCATCATGATCAAGACACGACACAGTCCG 79
QY 61 ttttctggtatctcccaagaatcactgctcagaataatgagtcacagaacttcag 120
80 TGTTCGTGATATCTCCAGAAATCAGCTTCAGAAATATGATGACACAGAACTTCAG 139
QY 121 coacgaccttcaactcaagcccttcgcaaaaattatctagaaaatgaaaatct 180
DB 140 CCAGACCTTTTCAACTCAAGCCCTTCGAAAATATATTCCTGAAAATGAAAATCT 199
QY 181 taggactacacagatcctglttgaaatagaccttcttcttgagttactctct 240
DB 200 TAGGACATATCCAGATCTGTTGGAATATGACCTTCTTTCGAGTATCTTCCTT 259
QY 241 tcaacttgtaaaacacatatacgaagttcccttataatcttcttcagagatccattct 300
DB 260 TCACCTTTTAAACCATATCCAGGTTCCCTTATATTCCTTTCAGGATATCATTCCT 319
QY 301 ggggctcgttltgtcaatlaatcttgagcccttcctaatgcaatgaaagaaaacaa 360
DB 320 GGGGCTCTGTTTGTCTATATTCGTGAGCCTTCCTAATTCAGATGAAAAGAAAACCA 379
QY 361 cagaacacctgataatattgagcgaataaagaatcctcttagtgcctgagagcaatag 420
DB 380 CAGAAACCTGTATATATGAGCCGAAATATGAAATTTCTTAGTCCCTGGAGCAATAG 439
QY 421 ctggaatccttcctcacatcttggttcaatctagatcaaacacacatttggttatt 480
DB 440 CTGGAATATATCTCTCACATTTGTTTCAATCTAGATCAACATCATTTGTTTAT 499
QY 481 ctcaacaaaatagtcagtgtaagcgtctactgctgtcttgaggaaatttgatacat 540
500 CTCACCAAAATAGTCAGTGTAGGCTGTACTGCTGTTCTTGGAATTTTGATTAAT 559
QY 541 tgaactcctcagaatattgaatattcaatctctgctctccttcacatttgagggtgc 600
DB 560 TGATATCTTTCAGCATATATGAAATATTCATTTCTGCTTCATATTTTGGGCTGC 619
QY 601 actcagaagatttgatgtgaaacaattgttgactagcactgtgaaataaagattg 660
DB 620 ACTCGAAGATTGTGATGTGACAAATGTTGTTGACTGACACTGTGAGATTAAGATG 679
QY 661 ttaaaataaaa 672
DB 680 TTAATAATATAA 691

RESULT 5
AX179781 603 bp DNA PAT 06-AUG-2001
LOCUS Sequence 1 from Patent W00146417.
DEFINITION AX179781
ACCESSION AX179781
VERSION AX179781.1 GI:15132144
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 603)
AUTHORS Walke,D.W. and Turner,C.A.
TITLE Human membrane proteins and polynucleotides encoding the same
JOURNAL having homology to cd20 proteins and ige receptors
Patent: WO 0146417-A 1 28-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Location/Qualifiers
source 1..603
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 162 a 125 c 104 g 212 t
ORIGIN

Query Match 89.2%; Score 603; DB 6; Length 603;
Best Local Similarity 100.0%; Pred. No. 1e-130;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 atggaatcaagacacgacacagtcggttcttgatcttccccaagaatcactgct 92
DB 1 ATGGAATTCAGACACCGACACAGTCGGTGTTCGTATTCCTCCAGAAATCACTGCT 60
QY 93 tgaataatgagtcacagaacttccagccagaccttcaactcaagaagcccttgcaa 152
DB 61 TCAGATATGAGTCACAGAACTTCAGCCAGACGACCTTTCAACACAAAGCCCTTGCAA 120
QY 153 aaattatctgcagaanaatgaaatccttaagagatcaccagatccttggaattatg 212
DB 121 AAATATTTGTCAGAAAATGAAAATCTTAGGACATATCCAGATCTCTGTTGGAATATG 180
QY 213 accttctcttgagatcttctcttccacttgtaaaacacatccaaagttccc 272
DB 181 ACCCTTCTTTCGAGTATCTTCCTTTCACCTTGTAAACCATATCCAGGTTTCCTC 240
QY 273 ttatattctcttcagagatccattctcgaggctctggttcttgatattccttgagcc 332
DB 241 TTTATATTTCTTTCAGGATATCATCTGAGGCTCTGTTTGTATTAATCTGAGGC 300
QY 333 ttcctaattgcagtgaaagaaaacacagaaatcctgataatattgagccgaataag 392
DB 301 TTCCTAATTCAGTGAAAAGAAAACACAGAACTTCGATATATTTGAGCCGGAATATG 360
QY 393 aatctcttagtgccctgagagcaatagctggaatcattctccacatttggttccatc 452
DB 361 AATCTTCTTAGTGCCCTGAGACCAATAGCTGGAATCATCTCTCAATTTGGTTCAATC 420
QY 453 ctgagtcaaaactacatttggtgtatctcaccaaaatagtcagtgtaagcgttact 512
DB 421 CTAGATCAAAACTACATTTTGTTATTCCTCACCAAAATAGTCAGTGTAGGCTGTTACT 480
QY 513 gtccgtctcttggaatttgatctacatgagatctcagcatatctgaaattacatt 572
DB 481 GTCGCTCTTCGGAATTTTGATTCATTTGATGATGATTCAGCATATTAATTAATTCAT 540
QY 573 tctctgctcttcacatatttgagggtgccaactcagaagatttgatgtgaaacaatgctgt 632
DB 541 TCTTCGCTTCTTCATATTTTGGGTGCGACCTCAGAGGATTTGATTTGTAACAATGTTGT 600
QY 633 tga 635
DB 601 TGA 603

RESULT 6
AX179787 450 bp DNA PAT 06-AUG-2001
LOCUS Sequence 7 from Patent W00146417.
DEFINITION AX179787
ACCESSION AX179787
VERSION AX179787.1 GI:15132147
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 450)
Walke, D.W. and Turner, C.A.
Human membrane proteins and polynucleotides encoding the same having homology to cd20 proteins and ige receptors
Patent: WO 0146417-A 7 28-JUN-2001;
Lexicon Genetics Incorporated (US)

FEATURES
source
location/Qualifiers

BASE COUNT 120 a 94 c 77 g 159 t
ORIGIN

Query Match 51.4%; Score 347.4; DB 6; Length 450;
Best Local Similarity 97.0%; Pred. No. 4e-71;
Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

33 atggatcgaagcagccagacagtcggtgttctgtatcttccccaagaatcgtct 92
|||||
1 atggattcagacacccacacacgctcggtgttctgtatcttccccaagaatcgtct 60
|||||
93 tcagaatcgaagtcacagaacttcagccagacacttcaactcaagaagcccttgcaa 152
|||||
61 tcagatatgagatccacacacacttccagccagacacttcaactcaagaagcccttgcaa 120
|||||
153 aaattatcgtcagaanaaaatgaataccttaggactatccagatccgttggaaattatg 212
|||||
121 aaattatcgtcagaanaaaatgaataccttaggactatccagatccgttggaaattatg 180
|||||
213 acccttcttggagttatcttcccttcaactgttaaacatcagaagttctcc 272
|||||
181 acccttcttggagttatcttcccttcaactgtttaaaccatcagaagttctcc 240
|||||
273 ttataatcttccagagatccatctcgtggtctgttctgtaataatcctgagcc 332
|||||
241 tttatattcttccagagatccatctcgtggtctgttctgtaataatcctgagcc 300
|||||
333 ttccatcgtcagtgagaaagaaacacagaaactcgtatataatgagccgaataatg 392
|||||
301 ttccatcgtcagtgagaaagaaacacagaaactcgtatataatgagccgaataatg 360
|||||
393 aatct 397
|||||
361 acttt 365

RESULT 7
AX179785 417 bp DNA PAT 06-AUG-2001
LOCUS AX179785
DEFINITION Sequence 5 from Patent WO0146417.
ACCESSION AX179785
VERSION AX179785.1 GI:15132146
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 417)
Walke, D.W. and Turner, C.A.
Human membrane proteins and polynucleotides encoding the same having homology to cd20 proteins and ige receptors
Patent: WO 0146417-A 5 28-JUN-2001;
Lexicon Genetics Incorporated (US)

FEATURES
source
location/Qualifiers

BASE COUNT 115 a 91 c 69 g 142 t
ORIGIN

Query Match 50.6%; Score 341.8; DB 6; Length 417;
Best Local Similarity 99.4%; Pred. No. 1e-70;
Matches 343; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

33 atggatcgaagcagccagacagtcggtgttctgtatcttccccaagaatcgtct 92
|||||
1 atggattcagacacccacacacgctcggtgttctgtatcttccccaagaatcgtct 60
|||||
93 tcagaatcgaagtcacagaacttcagccagacacttcaactcaagaagcccttgcaa 152
|||||
61 tcagatatgagatccacacacacttccagccagacacttcaactcaagaagcccttgcaa 120
|||||
153 aaattatcgtcagaanaaaatgaataccttaggactatccagatccgttggaaattatg 212
|||||
121 aaattatcgtcagaanaaaatgaataccttaggactatccagatccgttggaaattatg 180
|||||
213 acccttcttggagttatcttcccttcaactgtttaaaccatcagaagttctcc 272
|||||
181 acccttcttggagttatcttcccttcaactgtttaaaccatcagaagttctcc 240
|||||
273 ttataatcttccagagatccatctcgtggtctgttctgtaataatcctgagcc 332
|||||
241 tttatattcttccagagatccatctcgtggtctgttctgtaataatcctgagcc 300
|||||
333 ttccatcgtcagtgagaaagaaacacagaaactcgtatataatgagccgaataatg 377
|||||
301 ttccatcgtcagtgagaaagaaacacagaaactcgtatataatgagccgaataatg 345
|||||

RESULT 8
AX179783 246 bp DNA PAT 06-AUG-2001
LOCUS AX179783
DEFINITION Sequence 3 from Patent WO0146417.
ACCESSION AX179783
VERSION AX179783.1 GI:15132145
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 246)
Walke, D.W. and Turner, C.A.
Human membrane proteins and polynucleotides encoding the same having homology to cd20 proteins and ige receptors
Patent: WO 0146417-A 3 28-JUN-2001;
Lexicon Genetics Incorporated (US)

FEATURES
source
location/Qualifiers

BASE COUNT 60 a 45 c 47 g 94 t
ORIGIN

Query Match 36.4%; Score 246; DB 6; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.7e-47;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

390 atgaatcttcttagtcgcttgagagacatagcttgaaatcattctccatcattgtttc 449
|||||
1 atgaattcttcttagtcgcttgagagacatagcttgaaatcattctccatcattgtttc 60
|||||
450 atcctagatcaaaaactaacttctgtgtatcttcaacaaaatgacgtgtaagctgt 509
|||||
61 atccctagatcaaaaactaacttctgtgtatcttcaacaaaatgacgtgtaagctgt 120
|||||
510 actgctcgtcttgagaaatttgattacatgagactcagacttaataatc 569
|||||
121 actgctcgtcttgagaaatttgattacatgagactcagacttaataatc 180
|||||
570 attctcgtcttccatatttggtggtgacacacagagattgattgtaacaatgt 629
|||||

Db 181 ATTCTCTGCTTTCATTTTGGGGCTCCACTGACAGCATTTGATTGTGAACAATGT 240
Oy 630 tqtga 635
|||||
Db 241 TGTGTA 246

RESULT 9
AC027787
LOCUS AC027787
DEFINITION Homo sapiens chromosome 15 clone RP11-196E16 map 15, LOW-PASS
SEQUENCE SAMPLING:
AC027787
AC027787.2 GI:9645160
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 (bases 1 to 138097)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 15, clone RP11-196E16
unpublished
2 (bases 1 to 138097)
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavskiy, L., Boukhalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collins, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, J. S.,
Dodgson, S., Domingo, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Klein, J., Lecoq, K., Lamazares, R., Landers, T., Lehotzky, J.,
Levine, R., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrum, J., Meneus, L., Minova, T., Miranda, C., Mlenka, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tefaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 18, 2000 this sequence version replaced gi:7382629.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L9008
Center clone name: 196_E_16

* NOTE: This record contains 163 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 663: contig of 663 bp in length

* 664 763: gap of 100 bp
* 764 1464: contig of 701 bp in length
* 1465 1564: gap of 100 bp
* 1565 2268: contig of 704 bp in length
* 2269 2368: gap of 100 bp
* 2369 3042: contig of 674 bp in length
* 3043 3142: gap of 100 bp
* 3143 3820: contig of 678 bp in length
* 3821 3920: gap of 100 bp
* 3921 4626: contig of 706 bp in length
* 4627 4726: gap of 100 bp
* 4727 5423: contig of 697 bp in length
* 5424 5523: gap of 100 bp
* 5524 6234: contig of 711 bp in length
* 6235 6334: gap of 100 bp
* 6335 7040: contig of 706 bp in length
* 7041 7140: gap of 100 bp
* 7141 7837: contig of 697 bp in length
* 7838 7937: gap of 100 bp
* 7938 8632: contig of 695 bp in length
* 8633 8732: gap of 100 bp
* 8733 9430: contig of 698 bp in length
* 9431 9530: gap of 100 bp
* 9531 10214: contig of 684 bp in length
* 10215 10314: gap of 100 bp
* 10315 10980: contig of 666 bp in length
* 10981 11080: gap of 100 bp
* 11081 11783: contig of 703 bp in length
* 11784 11883: gap of 100 bp
* 11884 12582: contig of 699 bp in length
* 12583 12682: gap of 100 bp
* 12683 13389: contig of 707 bp in length
* 13390 13489: gap of 100 bp
* 13490 14186: contig of 697 bp in length
* 14187 14286: gap of 100 bp
* 14287 14996: contig of 710 bp in length
* 14997 15096: gap of 100 bp
* 15097 15796: contig of 700 bp in length
* 15797 15896: gap of 100 bp
* 15897 16583: contig of 687 bp in length
* 16584 16683: gap of 100 bp
* 16684 17376: contig of 683 bp in length
* 17377 17476: gap of 100 bp
* 17477 18178: contig of 702 bp in length
* 18179 18278: gap of 100 bp
* 18279 18962: contig of 684 bp in length
* 18963 19062: gap of 100 bp
* 19063 19673: contig of 611 bp in length
* 19674 19773: gap of 100 bp
* 19774 20436: contig of 663 bp in length
* 20437 20536: gap of 100 bp
* 20537 21233: contig of 697 bp in length
* 21234 21333: gap of 100 bp
* 21334 22042: contig of 709 bp in length
* 22043 22142: gap of 100 bp
* 22143 22808: contig of 666 bp in length
* 22809 22908: gap of 100 bp
* 22909 23595: contig of 687 bp in length
* 23596 23695: gap of 100 bp
* 23696 24391: contig of 696 bp in length
* 24392 24491: gap of 100 bp
* 24492 25190: contig of 699 bp in length
* 25191 25290: gap of 100 bp
* 25291 25981: contig of 691 bp in length
* 25982 26081: gap of 100 bp
* 26082 26765: contig of 684 bp in length
* 26766 26865: gap of 100 bp
* 26866 27563: contig of 698 bp in length
* 27564 27663: gap of 100 bp
* 27664 28377: contig of 714 bp in length
* 28378 28477: gap of 100 bp
* 28478 29184: contig of 707 bp in length
* 29185 29284: gap of 100 bp

```

* 29285 29983: contig of 699 bp in length
* 29984 30083: gap of 100 bp
* 30084 30775: contig of 692 bp in length
* 30776 30875: gap of 100 bp
* 30876 31665: contig of 690 bp in length
* 31666 31665: gap of 100 bp
* 31666 32360: contig of 695 bp in length
* 32361 32460: gap of 100 bp
* 32461 33157: contig of 697 bp in length
* 33158 33257: gap of 100 bp
* 33258 33951: contig of 694 bp in length
* 33952 34051: gap of 100 bp
* 34052 34735: contig of 684 bp in length
* 34736 34835: gap of 100 bp
* 34836 35483: contig of 648 bp in length
* 35484 35583: gap of 100 bp
* 35584 36308: contig of 725 bp in length
* 36309 36408: gap of 100 bp
* 36409 37108: contig of 700 bp in length
* 37109 37208: gap of 100 bp
* 37209 37892: contig of 684 bp in length
* 37893 37992: gap of 100 bp
* 37993 38672: contig of 680 bp in length
* 38673 38772: gap of 100 bp
* 38773 39475: contig of 703 bp in length
* 39476 39575: gap of 100 bp
* 39576 40290: contig of 715 bp in length
* 40291 40390: gap of 100 bp
* 40391 41087: contig of 697 bp in length
* 41088 41187: gap of 100 bp
* 41188 41691: contig of 704 bp in length
* 41692 41991: gap of 100 bp
* 41992 42679: contig of 688 bp in length
* 42680 42779: gap of 100 bp
* 42780 43500: contig of 721 bp in length
* 43501 43600: gap of 100 bp
* 43601 44295: contig of 695 bp in length
* 44296 44395: gap of 100 bp
* 44396 45088: contig of 693 bp in length
* 45089 45188: gap of 100 bp
* 45189 45880: contig of 692 bp in length
* 45881 45980: gap of 100 bp
* 45981 46683: contig of 703 bp in length
* 46684 46783: gap of 100 bp
* 46784 47481: contig of 698 bp in length
* 47482 47581: gap of 100 bp
* 47582 48290: contig of 709 bp in length
* 48291 48390: gap of 100 bp
* 48391 49077: contig of 687 bp in length
* 49078 49177: gap of 100 bp
* 49178 49879: contig of 702 bp in length
* 49880 49979: gap of 100 bp
* 49980 50684: contig of 705 bp in length
* 50685 50784: gap of 100 bp
* 50785 51507: contig of 723 bp in length
* 51508 51607: gap of 100 bp
* 51608 52320: contig of 713 bp in length
* 52321 52420: gap of 100 bp
* 52421 53133: contig of 713 bp in length
* 53134 53233: gap of 100 bp
* 53234 53933: contig of 700 bp in length
* 53934 54033: gap of 100 bp
* 54034 54703: contig of 670 bp in length

```

Query Match 27.4% Score 185.4; DB 2: Length 138097;
 Best Local Similarity 96.9%; Pred. No. 3.8e-33;
 Matches 189; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 tcatctccttcaaatatcacgacacatcatgatcatcaagcacgacacagtcg 60
 Db 109564 TCATCTCCCTTCAAAATATCACGACACATCATGATTCACGACACAGTCGCG 109623
 Qy 61 tgttctgtattctcctccagaatactcgtcttcagataatgagtcacagaacttcag 120

```

Db 109624 TGTTCGTGATTTCTCCAGAAATCATCTCTCAGAAATATGATGTCACAGACTTCAG 109683
Qy 121 cccagacctttaaactcaagcccttgcaaaatattgttagaaaatgaatc 180
Db 109684 CCAAGACCTTTCAACTCAAGCCCTTGCAAAATATTATTGCTAGCAAAATGAATCT 109743
Qy 181 taggactatccaga 195
Db 109744 TACGGGTAGTAGTA 109758

```

```

RESULT 10
AC027787/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
2 (bases 1 to 138097)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bede,F.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choquel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galegan,J., Gardyna,S., Glinde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hayes,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lechoczky,J.,
Levine,R., Lien,G., Liu,G., Locke,K., MacDonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McCurk,A., McKernan,K., McHeaters,R.,
Meidam,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,T., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Theodore,J., Titrrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 18, 2000 this sequence version replaced g1:7382629.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996:1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 19008
Center clone name: 196_E_16

```

* NOTE: This record contains 163 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.

* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
663: contig of 663 bp in length

664 763: gap of 100 bp
764 1464: contig of 701 bp in length
1465 1564: gap of 100 bp
1565 2268: contig of 704 bp in length
2269 2368: gap of 100 bp
2369 3042: contig of 674 bp in length
3043 3142: gap of 100 bp
3143 3820: contig of 678 bp in length
3821 3920: gap of 100 bp
3921 4626: contig of 706 bp in length
4627 4726: gap of 100 bp
4727 5423: contig of 697 bp in length
5424 5523: gap of 100 bp
5524 6234: contig of 711 bp in length
6235 6334: gap of 100 bp
6335 7040: contig of 706 bp in length
7041 7140: gap of 100 bp
7141 7837: contig of 697 bp in length
7838 7937: gap of 100 bp
7938 8632: contig of 695 bp in length
8633 8732: gap of 100 bp
8733 9430: contig of 698 bp in length
9431 9530: gap of 100 bp
9531 10214: contig of 684 bp in length
10215 10314: gap of 100 bp
10315 10980: contig of 666 bp in length
10981 11080: gap of 100 bp
11081 11783: contig of 703 bp in length
11784 11883: gap of 100 bp
11884 12582: contig of 699 bp in length
12583 12682: gap of 100 bp
12683 13389: contig of 707 bp in length
13390 13489: gap of 100 bp
13490 14186: contig of 697 bp in length
14187 14286: gap of 100 bp
14287 14996: contig of 710 bp in length
14997 15096: gap of 100 bp
15097 15796: contig of 700 bp in length
15797 15896: gap of 100 bp
15897 16583: contig of 687 bp in length
16584 16683: gap of 100 bp
16684 17376: contig of 693 bp in length
17377 17476: gap of 100 bp
17477 18178: contig of 702 bp in length
18179 18278: gap of 100 bp
18279 18962: contig of 684 bp in length
18963 19062: gap of 100 bp
19063 19673: contig of 611 bp in length
19674 19773: gap of 100 bp
19774 20436: contig of 663 bp in length
20437 20536: gap of 100 bp
20537 21233: contig of 697 bp in length
21234 21333: gap of 100 bp
21334 22042: contig of 709 bp in length
22043 22142: gap of 100 bp
22143 22808: contig of 666 bp in length
22809 22908: gap of 100 bp
22909 23595: contig of 687 bp in length
23596 23695: gap of 100 bp
23696 24391: contig of 696 bp in length
24392 24491: gap of 100 bp
24492 25190: contig of 699 bp in length
25191 25290: gap of 100 bp
25291 25981: contig of 691 bp in length
25982 26081: gap of 100 bp
26082 26765: contig of 684 bp in length
26766 26865: gap of 100 bp
26866 27563: contig of 698 bp in length

27564 27663: gap of 100 bp
27664 28377: contig of 714 bp in length
28378 28477: gap of 100 bp
28478 29184: contig of 707 bp in length
29185 29284: gap of 100 bp
29285 29983: contig of 699 bp in length
29984 30083: gap of 100 bp
30084 30775: contig of 692 bp in length
30776 30875: gap of 100 bp
30876 31565: contig of 690 bp in length
31566 31665: gap of 100 bp
31666 32360: contig of 695 bp in length
32361 32460: gap of 100 bp
32461 33157: contig of 697 bp in length
33158 33257: gap of 100 bp
33258 33951: contig of 694 bp in length
33952 34051: gap of 100 bp
34052 34735: contig of 684 bp in length
34736 34835: gap of 100 bp
34836 35483: contig of 648 bp in length
35484 35583: gap of 100 bp
35584 36308: contig of 725 bp in length
36309 36408: gap of 100 bp
36409 37108: contig of 700 bp in length
37109 37208: gap of 100 bp
37209 37892: contig of 684 bp in length
37893 37992: gap of 100 bp
37993 38672: contig of 680 bp in length
38673 38772: gap of 100 bp
38773 39475: contig of 703 bp in length
39476 39575: gap of 100 bp
39576 40290: contig of 715 bp in length
40291 40390: gap of 100 bp
40391 41087: contig of 697 bp in length
41088 41187: gap of 100 bp
41188 41891: contig of 704 bp in length
41892 41991: gap of 100 bp
41992 42679: contig of 688 bp in length
42680 42779: gap of 100 bp
42780 43500: contig of 721 bp in length
43501 43600: gap of 100 bp
43601 44295: contig of 695 bp in length
44296 44395: gap of 100 bp
44396 45088: contig of 693 bp in length
45089 45188: gap of 100 bp
45189 45880: contig of 692 bp in length
45881 45980: gap of 100 bp
45981 46683: contig of 703 bp in length
46684 46783: gap of 100 bp
46784 47481: contig of 698 bp in length
47482 47581: gap of 100 bp
47582 48290: contig of 709 bp in length
48291 48390: gap of 100 bp
48391 49077: contig of 687 bp in length
49078 49177: gap of 100 bp
49178 49879: contig of 702 bp in length
49880 49979: gap of 100 bp
49980 50684: contig of 705 bp in length
50685 50784: gap of 100 bp
50785 51507: contig of 723 bp in length
51508 51607: gap of 100 bp
51608 52320: contig of 713 bp in length
52321 52420: gap of 100 bp
52421 53133: contig of 713 bp in length
53134 53233: gap of 100 bp
53234 53933: contig of 700 bp in length
53934 54033: gap of 100 bp
54034 54703: contig of 670 bp in length

Query Match 27.4%; Score 185.4; DB 2; Length 138097;
Best Local Similarity 96.9%; Pred. No. 3.8e-33;
Matches 189; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 tcatctccttcaaatatcaccgacacatcatgatgattcaagcaccgacacagtcg 60
    |||||||
DB 112577 TCATCTCCTTTCAAAATTATACACGACATCATGATTCAGACACCGACACAGTCGCG 112518
QY 61 tttcttggtatltccctccagaatacactgcttcagaataatgctccacagaacttcag 120
    |||||||
DB 112517 TCTTTCTGTAATTTCTCCAGAAATCAGCTTCGAAATATGATGATCCACAGAACTTTGAG 112458
QY 121 ccagacacttcaactcaagcccttcgaaataatttctgtagaataatgaataatct 180
    |||||||
DB 112457 CCACGACCTTTTCAACTCAAGCCCTTCGAAAAATATTCTGTAGAAAAATGAATATCT 112398
QY 181 taggactatccaga 195
    |||||
DB 112397 TAGGGTAGTAAGA 112383

RESULT 11
AC015840 147788 bp DNA HTG 27-APR-2000
LOCUS Homo sapiens clone RP11-24D1, WORKING DRAFT SEQUENCE, 5 unordered
DEFINITION
AC015840
AC015840.2 GI:7657730
VERSION
HTG: HTGS_PHASE1, HTGS_DRAFT.
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 147788)
Britten,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL
Unpublished
AUTHORS
2 (bases 1 to 147788)
Britten,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckert,L., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collimore,A.,
Cooke,P., Dearliano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,N.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGuff,A., McKernan,K., McDonald,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,J.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Sudmanian,A., Talamas,J.,
Testa,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 27, 2000 this sequence version replaced g1:6446828.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 14434
Center clone name: 24.D.1
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 136347 bases at least Q40
Consensus quality: 143600 bases at least Q30
Consensus quality: 146212 bases at least Q20
Insert size: 205000; agarose-fp
Insert size: 147388; sum-of-ctnigs
Quality coverage: 5.2 in Q20 bases; agarose-fp

```

```

Quality coverage: 7.1 in Q20 bases; sum-of-ctnigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 6594: contig of 6594 bp in length
* 6595 6694: gap of 100 bp
* 6695 16187: contig of 9493 bp in length
* 16188 16287: gap of 100 bp
* 16288 25089: contig of 8802 bp in length
* 25090 25189: gap of 100 bp
* 25190 63923: contig of 38734 bp in length
* 63924 64023: gap of 100 bp
* 64024 147788: contig of 83765 bp in length.
FEATURES
Source
1..147788
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-24D1"
/clone_lib="RPCT-11 Human Male BAC"
1..6594
/note="assembly-fragment"
clone_end:17
vector_side:left"
6695..16187
/note="assembly-fragment"
16288..25089
/note="assembly-fragment"
clone_end:SP6
vector_side:right"
25190..63923
/note="assembly-fragment"
64024..147788
/note="assembly-fragment"
BASE COUNT 45466 a 29482 c 28861 g 43569 t 410 others
ORIGIN
Query Match 27.4%; Score 185.4; DB 2; Length 147788;
Best Local Similarity 96.9%; Pred. No. 3.8e-33;
Matches 189; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 tcatctccttcaaatatcaccgacacatcatgatgattcaagcaccgacacagtcg 60
    |||||||
DB 128015 TCATCTCCTTTCAAAATTATACACGACATCATGATTCAGACACCGACACAGTCGCG 128074
QY 61 tttcttggtatltccctccagaatacactgcttcagaataatgctccacagaacttcag 120
    |||||||
DB 128075 TCTTTCTGTAATTTCTCCAGAAATCAGCTTCGAAATATGATGATCCACAGAACTTTGAG 128134
QY 121 ccagacacttcaactcaagcccttcgaaataatttctgtagaataatgaataatct 180
    |||||||
DB 128135 CCACGACCTTTTCAACTCAAGCCCTTCGAAAAATATTCTGTAGAAAAATGAATATCT 128194
QY 181 taggactatccaga 195
    |||||
DB 128195 TAGGGTAGTAAGA 128209

RESULT 12
AP003127 161039 bp DNA HTG 30-JAN-2001
LOCUS Homo sapiens chromosome 11 clone RP11-7106 map 11q, WORKING DRAFT
DEFINITION
SEQUENCE, 15 unordered pieces.
ACCESSION
AP003127
AP003127.1 GI:12597183
VERSION
HTG: HTGS_PHASE1, HTGS_DRAFT.
KEYWORDS
Homo sapiens DNA, clone:RP11-7106.

```

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 161039)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Matanabe,H. and Sakaki,Y.
 TITLE Homo sapiens 161,039 genomic DNA of 11q
 JOURNAL Published Only in Database (2001) In press
 REFERENCE 2 (bases 1 to 161039)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Matanabe,H. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (26-JAN-2001) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 COMMENT

Center: RIKEN Genomic Sciences Center(GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@gsc.riken.go.jp

Project Information

Center project name: Humdraft11

Center clone name: Rp11-71066

Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 157456 bases at least Q40

Consensus quality: 159457 bases at least Q30

Insert size: 159639, sum-of-ctrls

Quality coverage: 8.61x in Q20 bases; sum-of-ctrls

NOTE: This is a 'working draft' sequence. It currently consists of

15 contigs. The true order of the pieces is not known and their

order in this sequence record is arbitrary. Gaps between the

contigs are represented as runs 'N', but the exact sizes of the gaps

are unknown. This record will be updated with the finished sequence

as soon as it is available and the accession number will be

preserved

1 42322 contig of 42322 bp in length

42423 63418 contig of 20996 bp in length

63519 77969 contig of 14451 bp in length

78070 93182 contig of 15113 bp in length

93283 104093 contig of 10811 bp in length

104194 115702 contig of 11509 bp in length

115803 125183 contig of 9381 bp in length

125284 132558 contig of 7275 bp in length

132659 137219 contig of 4561 bp in length

137320 143276 contig of 5957 bp in length

143377 148922 contig of 5546 bp in length

149023 154159 contig of 5137 bp in length

154260 157686 contig of 3427 bp in length

157787 159234 contig of 1448 bp in length

159335 161039 contig of 1705 bp in length

NOTE: This is a 'working draft' sequence. It currently

consists of 15 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of 'N', but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 42322: contig of 42322 bp in length

42323 42423: gap of 100 bp

42424 63418: contig of 20996 bp in length

63419 63519: gap of 100 bp

63520 77969: contig of 14451 bp in length

77970 78069: gap of 100 bp

78070 93182: contig of 15113 bp in length

93183 93282: gap of 100 bp

93283 104093: contig of 10811 bp in length

104094 104193: gap of 100 bp

104194 115702: contig of 11509 bp in length

115703 115802: gap of 100 bp

115803 125183: contig of 9381 bp in length

125184 125283: gap of 100 bp

125284 132558: contig of 7275 bp in length

132559 132658: gap of 100 bp

132659 137219: contig of 4561 bp in length

137220 137319: gap of 100 bp

137320 143276: contig of 5957 bp in length

143277 143376: gap of 100 bp

143377 148922: contig of 5546 bp in length

148923 149022: gap of 100 bp

149023 154159: contig of 5137 bp in length

154160 154259: gap of 100 bp

154260 157686: contig of 3427 bp in length

157687 157786: gap of 100 bp

157787 159234: contig of 1448 bp in length

159235 159334: gap of 100 bp

159335 161039: contig of 1705 bp in length.

Location/Qualifiers

1. 161039

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="11"

/map="11q"

/clone="Rp11-71066"

1. 42322

/note="assembly_fragment"

42423. 63418

/note="assembly_fragment"

63519. 77969

/note="assembly_fragment"

78070. 93182

/note="assembly_fragment"

93283. 104093

/note="assembly_fragment"

104194. 115702

/note="assembly_fragment"

115803. 125183

/note="assembly_fragment"

125284. 132558

/note="assembly_fragment"

132659. 137219

/note="assembly_fragment"

137320. 143276

/note="assembly_fragment"

143377. 148922

/note="assembly_fragment"

148923. 149022

/note="assembly_fragment"

149023. 154159

/note="assembly_fragment"

154260. 157686

/note="assembly_fragment"

157787. 159234

/note="assembly_fragment"

159335. 161039

/note="assembly_fragment"

BASE COUNT 49155 a 31166 c 31523 g 47795 t 1400 others

ORIGIN

Query Match 27.4%; Score 185.4; DB 2; Length 161039;

Best Local Similarity 96.9%; Pred. No. 3.8e-33;

Matches 189; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

DB 149650 TCATCTCTTCAATTAATTCACGACACATCAGATTCAGAACCCACACAGTCGG 149591

DB 149590 TGTTCGTGATTTCTCCCAAAATCACTGCTTCAGATATGATGCCACAGACTTTCAG 149531

Best Local Similarity 96.9%; Pred. No. 3.9e-33;
Matches 189; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 |tattccctttaaatttaccgacacatgatgattcaagaccgacacagtcg 60
Db 144133 TCATCTCCCTTAAATATACCGACCATCATGATTCAGACCGACACAGCTTCG 144073
QY 61 tgttcttgatcttcctccagaatcactgcttcagatagtcacagacttcag 120
Db 144072 TGTTCCTGGTATTTCTCCAGAAATCAGCTTCAGAAATATGATTCACAGAACTTCAG 144013
QY 121 ccagacaccttcaactcaagcccttcgaaaattattgctagaaaatgaaatct 180
Db 144012 CCACGACCTTTCAACTCAAGCCCTTCAGAAATATTTATTCGTAGAAAATGAAATCT 143953
QY 181 taggacatctcaga 195
Db 143952 TAGGGGTAAAGTAAGA 143938

RESULT 15

LOCUS AP003127 161039 bp DNA HTG 30-JAN-2001
DEFINITION Homo sapiens chromosome 11 clone RP11-710G6 map 11q, WORKING DRAFT
ACCESSION AP003127
VERSION AP003127.1 GI:12597183
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 161039)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 161,039 genomic DNA of 11q
Published Only in Database (2001) In press
2 (bases 1 to 161039)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE

Journal Submitted (26-JAN-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)

AUTHORS

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

JOURNAL

Journal Submitted (26-JAN-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT

Center: RIKEN Genomic Sciences Center(GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@gsc.riken.go.jp

Project Information

Center project name: HumDraft11

Center clone name: RP11-710G6

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 157456 bases at least Q40

Consensus quality: 158960 bases at least Q30

Insert size: 159639; sum-of-contigs

Quality coverage: 8.61x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
15 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved

42423 1 42322 contig of 42322 bp in length
63418 contig of 20996 bp in length

FEATURES

Source

63519 77969 contig of 14451 bp in length
78070 93182 contig of 15113 bp in length
93283 104093 contig of 10811 bp in length
104194 115702 contig of 11509 bp in length
115803 125183 contig of 9381 bp in length
125284 132558 contig of 7275 bp in length
132659 137219 contig of 4561 bp in length
137320 143276 contig of 5546 bp in length
143377 148922 contig of 5137 bp in length
149023 154159 contig of 5137 bp in length
154260 157686 contig of 3427 bp in length
157787 159234 contig of 1448 bp in length
159235 159334 gap of 100 bp
159335 161039 contig of 1705 bp in length.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 42322 contig of 42322 bp in length
42323 42422 gap of 100 bp
42423 63418 contig of 20996 bp in length
63419 63518 gap of 100 bp
63519 77969 contig of 14451 bp in length
77970 78069 gap of 100 bp
78070 93182 contig of 15113 bp in length
93183 93282 gap of 100 bp
93283 104093 contig of 10811 bp in length
104094 104193 gap of 100 bp
104194 115702 contig of 11509 bp in length
115703 115802 gap of 100 bp
115803 125183 contig of 9381 bp in length
125184 125283 gap of 100 bp
125284 132558 contig of 7275 bp in length
132559 132658 gap of 100 bp
132659 137219 contig of 4561 bp in length
137220 137319 gap of 100 bp
137320 143276 contig of 5546 bp in length
143277 143376 gap of 100 bp
143377 148922 contig of 5137 bp in length
148923 149022 gap of 100 bp
149023 154159 contig of 5137 bp in length
154160 154259 gap of 100 bp
154260 157686 contig of 3427 bp in length
157687 157786 gap of 100 bp
157787 159234 contig of 1448 bp in length
159235 159334 gap of 100 bp
159335 161039 contig of 1705 bp in length.

FEATURES

Source

1. 161039
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-710G6"
1. 42322
/note="assembly-fragment"
42423. 63418
/note="assembly-fragment"
63519. 77969
/note="assembly-fragment"
78070. 93182
/note="assembly-fragment"
93283. 104093
/note="assembly-fragment"
104194. 115702
/note="assembly-fragment"
115803. 125183
/note="assembly-fragment"
125284. 132558
/note="assembly-fragment"

```

misc_feature      132659..137219
                  /note="assembly_fragment clone_end:SP6 vector_side:left"
misc_feature      137320..143276
                  /note="assembly_fragment"
misc_feature      143377..148922
                  /note="assembly_fragment clone_end:T7 vector_side:left"
misc_feature      149023..154159
                  /note="assembly_fragment"
misc_feature      154260..157686
                  /note="assembly_fragment"
misc_feature      157787..159234
                  /note="assembly_fragment"
misc_feature      159335..161039
                  /note="assembly_fragment"
BASE COUNT      49155 a 31166 c 31523 g 47795 t 1400 others
ORIGIN

```

```

Query Match      23.2%; Score 157; DB 2; Length 161039;
Best Local Similarity 91.7%; Pred. No. 1.6e-26;
Matches 166; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY      368      tctgataataggacgcaataatgaatctctcttagtgcctgagagcaatagctggaat 427
          |||||||
Db      45928      TCAGATAATATTGACCGCAATATATTTCTTAGTGCCCTGGAGCAATAGCTGAAT 45987

OY      428      catctccctcacattgtttcatctctagatcaaaactacatttggttatctcacca 487
          |||||||
Db      45988      CATCTCCTCACATTTGGTTTCATCTAGATCAAAACTACATTTGGTTATTCACACCA 46047

OY      488      aaatagtcagtgtaaggctgtctactgtctctgttcttgggaatttgattacattgatgac 547
          |||||||
Db      46048      AAATAGTCAGTGTAGGCTGTACTGTCTGTTCTTGGTAGTAGTATTCATTATAGAG 46107

OY      548      t 548
          |
Db      46108      T 46108

```

Search completed: November 28, 2001, 23:52:04
 Job time: 8367 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 00:23:54 ; Search time 204.6 Seconds

(without alignments)
2832,609 Million cell updates/sec

Title: US-09-735-712-9

Perfect score: 676

Sequence: 1 tcatctcccttcaattatc.....tjgttcaataataaaaaa 676

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N.Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	676	100.0	676	AA504283	DNA sequence encod
2	668	98.8	747	AAH64741	Human secreted pro
3	603	89.2	603	AA504279	Novel human membra
4	347.4	51.4	450	AA504282	Novel human membra
5	341.8	50.6	417	AA504281	Novel human membra
6	312.8	46.3	372	AA503080	Human secreted pro
7	246	36.4	246	AA504280	Novel human membra
8	190.4	28.2	468	AAH64745	Human secreted pro
9	108.8	16.1	936	AA58252	Oligonucleotide D1
10	108.8	16.1	936	AA58252	Oligonucleotide D1
11	108.8	16.1	936	AA58254	Oligonucleotide D1

c	12	108.8	16.1	936	22	AA58254	Oligonucleotide D1
c	13	108.8	16.1	936	22	AA58257	Oligonucleotide D1
c	14	108.8	16.1	936	22	AA58259	Oligonucleotide D1
c	15	108.8	16.1	936	22	AA58259	Oligonucleotide D2
c	16	108.8	16.1	936	22	AA58259	Oligonucleotide D2
c	17	108.8	16.1	936	22	AA58262	Oligonucleotide D2
c	18	108.8	16.1	936	22	AA58262	Oligonucleotide D2
c	19	108.8	16.1	936	22	AA58262	Oligonucleotide D2
c	20	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	21	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	22	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	23	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	24	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	25	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	26	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	27	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	28	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	29	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	30	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	31	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	32	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	33	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	34	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	35	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	36	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	37	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	38	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	39	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	40	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	41	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	42	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	43	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	44	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	45	108.8	16.1	936	22	AA58255	Oligonucleotide D1

ALIGNMENTS

RESULT	1
AA504283	standard; DNA; 676 BP.
ID	AA504283;
AC	AA504283;
XX	26-SEP-2001 (first entry)
DT	
XX	
DE	DNA sequence encoding novel human membrane protein.
XX	
KW	Human; membrane protein; membrane receptor; Ige receptor; CD20;
KW	physiological disorder; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200146417-A2.
XX	
PD	28-JUN-2001.
XX	
PF	12-DEC-2000; 2000WO-US33742.
XX	
PR	22-DEC-1999; 99US-0171567.
XX	
RA	(LEXI-) LEXICON GENETICS INC.
XX	
PI	Walke DW, Turner CA;
XX	
DR	WPI; 2001-408646/43.
XX	
PT	Polynucleotide encoding novel human membrane protein, useful for
PT	identifying agonist, antagonist or modifiers or for producing
PT	antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT	applications
XX	

PS Disclosure: Page 31-32; 32pp; English.

XX The present sequence encoding for a novel human membrane protein
CC includes the flanking 5'- and 3'- sequences. Four amino acid sequences
CC for novel human membrane proteins (AAU01210-AAU01213) are given in the
CC present invention. These membrane proteins share structural similarity
CC with membrane receptors such as the IGF receptor and mammalian CD20.
CC The novel human membrane proteins are useful for identifying agonist,
CC antagonist and modulators of the membrane proteins, and for producing
CC antibodies specific to the membrane proteins. The membrane proteins can
CC be used for diagnosis, drug screening, pharmacogenomic applications,
CC clinical trial monitoring and the treatment of physiological disorders
CC and diseases. The polynucleotides encoding the membrane proteins can be
CC used to generate PCR primers or probes to identify mutations associated
CC with a particular disease.
XX

Sequence 676 BP; 193 A; 140 C; 112 G; 231 T; 0 other;

Query Match 100.0%; Score 676; DB 22; Length 676;
Best Local Similarity 100.0%; Pred. No. 3,7e-174;
Matches 676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tcattccttcaaatctccagcaccatcatgattcaagcaccgacacagtcg 60
Db 1 tcattccttcaaatctccagcaccatcatgattcaagcaccgacacagtcg 60
QY 61 ttttctggtatttccctccagaatcagcttcagaatgagtcacacagattcag 120
Db 61 ttttctggtatttccctccagaatcagcttcagaatgagtcacacagattcag 120
QY 121 ccaagacctttcaactaaagcccttgcaaaaattttgtcagaaatgaataatc 180
Db 121 ccaagacctttcaactaaagcccttgcaaaaattttgtcagaaatgaataatc 180
QY 181 taggagacatccagatcctgtttgaattagacatttcttttgagattatctt 240
Db 181 taggagacatccagatcctgtttgaattagacatttcttttgagattatctt 240
QY 241 tcacctgttaaaacatatacagaagttcccttataattcttccagatataatc 300
Db 241 tcacctgttaaaacatatacagaagttcccttataattcttccagatataatc 300
QY 301 ggggctcgtttgttcttaattcttcagagcttcccttaattgcagtgaaagaaac 360
Db 301 ggggctcgtttgttcttaattcttcagagcttcccttaattgcagtgaaagaaac 360
QY 361 cagaacctgtataatattgagccgaataatgaatcttctagtgcccttgagcaatag 420
Db 361 cagaacctgtataatattgagccgaataatgaatcttctagtgcccttgagcaatag 420
QY 421 ctggaatcattctcctcattgttctcattagatcaaaactactgtgtatt 480
Db 421 ctggaatcattctcctcattgttctcattagatcaaaactactgtgtatt 480
QY 481 ctcccaaaatagtcagtgtaaggctgtactgtcgttcttggaatttgattat 540
Db 481 ctcccaaaatagtcagtgtaaggctgtactgtcgttcttggaatttgattat 540
QY 541 tgaatgacttcagatattgaattatcatttctcgtccttccaaatttggtg 600
Db 541 tgaatgacttcagatattgaattatcatttctcgtccttccaaatttggtg 600
QY 601 actcagaggtattgattgtgaaacatggttctgactagcacttgagaataagattg 660
Db 601 actcagaggtattgattgtgaaacatggttctgactagcacttgagaataagattg 660
QY 661 ttaataaaaaaaa 676
Db 661 ttaataaaaaaaa 676

RESULT 2

AAH64741
ID AAH64741 standard; cDNA; 747 BP.

XX AAH64741;

AC 11-SEP-2001 (first entry)

DT Human secreted protein cDNA, SEQ ID NO: 17.

DE Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;

KW GENSET; ss.

OS Homo sapiens.

PN WO200142451-A2.

PD 14-JUN-2001.

PF 07-DEC-2000; 2000WO-IB01938.

PR 08-DEC-1999; 99US-0169629.

PR 06-MAR-2000; 2000US-0187470.

XX (GENSET) GENSET.

XX Dumas Milne Edwards J, Bougueleret L, Jobert S;

XX WPI; 2001-367870/38.

XX P-PSDB; AAG89138.

PS Claim 7; Page 576-577; 921pp; English.

CC The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased
CC GENSET gene expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of GENSET or by supplementing
CC the patient's own production of GENSET polypeptides. Conversely,
CC antisense nucleic acid molecules may be administered to down regulate
CC GENSET expression by binding with the cells' own genes and preventing
CC their expression. The sense and antisense nucleic acids may also be
CC used as DNA probes in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acid sequences in samples, and hence to
CC determine which patients may be in need of restorative therapy.
CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET nucleic acid of the invention.

Sequence 747 BP; 240 A; 145 C; 121 G; 241 T; 0 other;

Query Match 98.8%; Score 668; DB 22; Length 747;
Best Local Similarity 99.3%; Pred. No. 5,7e-172;
Matches 671; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 tcattccttcaaatctccagcaccatcatgattcaagcaccgacacagtcg 60
Db 1 tcattccttcaaatctccagcaccatcatgattcaagcaccgacacagtcg 60
QY 61 ttttctggtatttccctccagaatcagcttcagaatgagtcacacagattcag 120
Db 61 ttttctggtatttccctccagaatcagcttcagaatgagtcacacagattcag 120
QY 101 ttttctggtatttccctccagaatcagcttcagaatgagtcacacagattcag 160
Db 101 ttttctggtatttccctccagaatcagcttcagaatgagtcacacagattcag 160
QY 121 ccaagacctttcaactaaagcccttgcaaaaattttgtcagaaatgaataatc 180
Db 121 ccaagacctttcaactaaagcccttgcaaaaattttgtcagaaatgaataatc 180
QY 161 ccaagacctttcaactaaagcccttgcaaaaattttgtcagaaatgaataatc 220
Db 161 ccaagacctttcaactaaagcccttgcaaaaattttgtcagaaatgaataatc 220

RESULT 4
AAS04282
ID AAS04282 standard; CDNA: 450 BP.
XX
AC AAS04282;
XX
DT 26-SEP-2001 (first entry)
XX
DE Novel human membrane protein #4 CDNA sequence.
XX
KW Human: membrane protein; membrane receptor; IGE receptor; CD20;
KW physiological disorder; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..450
FT /tag= a
FT /product= "Human membrane protein #4"

W0200146417-A2.
XX
PD 28-JUN-2001.
XX
PE 12-DEC-2000; 2000WO-US33742.
XX
PR 22-DEC-1999; 99US-0171567.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Turner CA;
XX
DR WPI: 2001-408646/43.
DR P-PSDB; AAU01213.
XX
PT Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications
XX
PS Disclosure; Page 31; 32pp; English.

XX
XX The present sequence encodes for novel human membrane protein #4.
CC Human membrane protein #4 is 1 of 4 human membrane proteins
CC (AAU01210-AAU01213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the IGE receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.
XX
XX Sequence 450 BP; 120 A; 94 C; 77 G; 159 T; 0 other;

Query Match 51.4%; Score 347.4; DB 22; Length 450;
Best Local Similarity 97.0%; Pred. No. 5.9e-85;
Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
OY 33 atgattcaagcagccagcagctggttctgtgtattcttcctcagaatactgct 92
DB 1 atgattcaagcagccagcagctggttctgtgtattcttcctcagaatactgct 60
OY 93 tcagaatagtgctcagcagacttcagcagaccccttcaacicaagcccttgcaa 152
DB 61 tcagaatagtgctcagcagacttcagcagaccccttcaacicaagcccttgcaa 120
OY 153 aaattattgctagaanaaatacttaggactatccagatcctgtttggaattatg 212

DB 121 aaattattgctagaanaaatacttaggactatccagatcctgtttggaattatg 180
OY 213 acccttcttttgaggtatctctcttcacacctgtttaaacatcaccaagtttccc 272
DB 181 acccttcttttgaggtatctctcttcacacctgtttaaacatcaccaagtttccc 240
OY 273 ttatattcttcagagatccatctcgggctctgtttgttcatatccttgagcc 332
DB 241 ttatattcttcagagatccatctcgggctctgtttgttcatatccttgagcc 300
OY 333 ttccaatgcaagtgaaaagaaacccagaaatctcgataatattgagccgaataatg 392
DB 301 ttccaatgcaagtgaaaagaaacccagaaatctcgataatattgatacttgatg 360
OY 393 aatct 397
DB 361 acttt 365

RESULT 5
AAS04281
ID AAS04281 standard; CDNA: 417 BP.
XX
AC AAS04281;
XX
DT 26-SEP-2001 (first entry)
XX
DE Novel human membrane protein #3 CDNA sequence.
XX
KW Human: membrane protein; membrane receptor; IGE receptor; CD20;
KW physiological disorder; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..417
FT /tag= a
FT /product= "Human membrane protein #3"

W0200146417-A2.
XX
PD 28-JUN-2001.
XX
PE 12-DEC-2000; 2000WO-US33742.
XX
PR 22-DEC-1999; 99US-0171567.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Turner CA;
XX
DR WPI: 2001-408646/43.
DR P-PSDB; AAU01212.

XX
XX Polynucleotide encoding novel human membrane protein, useful for
XX identifying agonist, antagonist or modifiers or for producing
XX antibodies useful in therapeutic, diagnostic and pharmacogenomic
XX applications
XX
PS Disclosure; Page 30; 32pp; English.
XX
XX The present sequence encodes for novel human membrane protein #3.
CC Human membrane protein #3 is 1 of 4 human membrane proteins
CC (AAU01210-AAU01213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the IGE receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate

CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within t

XX 12-DEC-2000; 2000WO-US33742.
PF
XX
PR 22-DEC-1999; 9905-0171567.

PR 22-DEC-1999; 99US-0171567.
XX

PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Turner CA;
 XX
 DR WPI: 2001-408646/43.
 P-PSDB: AAU01211.
 XX
 PT Polynucleotide encoding novel human membrane protein, useful for
 PT identifying agonist, antagonist or modifiers or for producing
 PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
 PT applications
 XX
 PS Disclosure: Page 30; 32pp; English.
 XX
 CC The present sequence encodes for novel human membrane protein #2.
 CC Human membrane protein #2 is 1 of 4 human membrane proteins
 CC (AAU01210-AAU01213) given in the present invention. These membrane
 CC proteins share structural similarity with membrane receptors such as
 CC the IGF receptor and mammalian CD20. The novel human membrane proteins
 CC are useful for identifying agonists, antagonists and modulators of the
 CC membrane proteins, and for producing antibodies specific to the
 CC membrane proteins. The membrane proteins can be used for diagnosis,
 CC drug screening, pharmacogenomic applications, clinical trial monitoring
 CC and the treatment of physiological disorders and diseases. The
 CC polynucleotides encoding the membrane proteins can be used to generate
 CC PCR primers or probes to identify mutations associated with a particular
 CC disease.
 CC
 XX
 XX Sequence 246 BP; 60 A; 45 C; 47 G; 94 T; 0 other;
 SQ

Query Match 36.4%; Score 246; DB 22; Length 246;
 Best Local Similarity 100.0%; Pred. No. 1.7e-57;
 Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 390 atgaattcttagtgagcctgagagcaatagtgatcattcctcattgtttc 449
 |||
 Db 1 atgaattcttagtgagcctgagagcaatagtgatcattcctcattgtttc 60

QY 450 atcctagatcaaaaatacatctgtgtatctccacaaaatagtcagtgtaagctgt 509
 |||
 Db 61 atcctagatcaaaaatacatctgtgtatctccacaaaatagtcagtgtaagctgt 120

QY 510 actgtccgtctcttgaggaatttgatcatgagcttcagcattatgaattatc 569
 |||
 Db 121 actgtccgtctcttgaggaatttgatcatgagcttcagcattatgaattatc 180

QY 570 attctcgtccttctcaatttgggtggtccactcagaggaattgtgattgaacaatgt 629
 |||
 Db 181 attctcgtccttctcaatttgggtggtccactcagaggaattgtgattgaacaatgt 240

QY 630 tgttga 635
 |||
 Db 241 tgttga 246

RESULT 8
 AAH64745
 ID AAH64745 standard; cDNA; 468 BP.
 XX
 AC AAH64745;
 XX
 DT 11-SEP-2001 (first entry)
 XX
 DE Human secreted protein cDNA, SEQ ID NO: 21.
 XX
 KW Human: secreted protein; gene therapy; vaccine; treatment; diagnosis;
 KM GENSER; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200142451-A2.
 XX

PD 14-JUN-2001.
 XX
 XX 07-DEC-2000; 2000WO-IB01938.
 XX
 XX 08-DEC-1999; 99US-0169629.
 PR 06-MAR-2000; 2000US-0187470.
 XX
 XX (GENT) GENSER.
 PA
 PI Dumas Milne Edwards J, Bougueleret L, Jobert S;
 XX
 XX WPI: 2001-367870/38.
 DR P-PSDB: AAG89142.
 XX
 XX Full length GENSER human nucleic acids encoding potentially secreted
 PT proteins, useful in gene therapy and vaccination against a variety of
 PT diseases, and for diagnosis of those diseases -
 XX
 PS Claim 7; Page 581; 921pp; English.
 XX
 CC The invention relates to full length GENSER human nucleic acids encoding
 CC potentially secreted proteins. The nucleic acids and the polypeptides
 CC they encode may be used in the prevention, treatment and diagnosis of
 CC diseases associated with inappropriate GENSER gene expression. For
 CC example, they be used to treat disorders associated with decreased
 CC GENSER gene expression by rectifying mutations or deletions in a
 CC patient's genome that affect the activity of GENSER or by supplementing
 CC the patients own production of GENSER polypeptides. Conversely,
 CC antisense nucleic acid molecules may be administered to down regulate
 CC GENSER expression by binding with the cells' own genes and preventing
 CC their expression. The sense and antisense nucleic acids may also be
 CC used as DNA probes in diagnostic assays to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples, and hence to
 CC determine which patients may be in need of restorative therapy.
 CC The GENSER polypeptides may be used as antigens in the production of
 CC antibodies and in assays to identify modulators (agonists and
 CC antagonists) of GENSER polypeptide expression and activity. The
 CC present sequence is a GENSER nucleic acid of the invention.
 CC
 XX
 XX Sequence 468 BP; 155 A; 94 C; 81 G; 138 T; 0 other;
 SQ

Query Match 28.2%; Score 190.4; DB 22; Length 468;
 Best Local Similarity 77.7%; Pred. No. 2.7e-42;
 Matches 230; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 1 tcatctccttcaaatatcaacgacacacatcagatcagaacgacacagtcg 60
 |||
 Db 44 tcatctccttcaaatatcaacgacacacatcagatcagaacgacacagtcg 103

QY 61 tttctgtatattctccagaatacactgtctcagaatagtcacagaacttcag 120
 |||
 Db 104 tttctgtatattctccagaatacactgtctcagaatagtcacagaacttcag 163

QY 121 ccaagacctttcaactaaagcccttgcaaaaataattgtcagaaaatgaaatc 180
 |||
 Db 164 ccaagacctttcaactaaagcccttgcaaaaataattgtcagaaaatgaaatc 223

QY 181 taggagctacagatcctgtttggaattagaccttcttgggtgtatctctt 240
 |||
 Db 224 taggagatccatctcggggtctgtttgttcatcaattctggagaccttccatgc 283

QY 241 tcacctgttaaacacatatacgaagttcccttataattcttcagagatatca 296
 |||
 Db 284 agtgaagaagaaacacagaaactctgggaattttgattatcatgtatgacttca 339

RESULT 9
 AAF58252
 ID AAF58252 standard; DNA; 936 BP.
 XX
 AC AAF58252;
 XX

DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1835.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
RW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
FR 17-MAR-2000; 2000US-0190259.
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
UMek RM;
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
XS Example 6; Page 127; 15pp; English.

The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.

Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

Query Match	16.1%	Score 108.8	DB 22	Length 936
Best Local Similarity	0.7%	Pred. No. 5	1e-20	
Matches	5	Conservative	418	Mismatches 245; Indels 0; Gaps 0
QY	9	tttcaaatatcacgcacacatcatcattgattcaagcaccgcacacagtcggtttctc	68	
Db	59	tttcaaatatcacgcacacacatcatcattgattcaagcaccgcacacagtcggtttctc	118	
QY	69	gtatttcctccagaatactactgcttcagaatactgtgtcacagaactttcaccaagacc	128	
Db	119	gtatttcctccagaatactactgcttcagaatactgtgtcacagaactttcaccaagacc	178	
QY	129	ttttcaactcaaaagcccttcgaaaattattgtctagaaaaatgaaaactcttaggact	188	
Db	179	ttttcaactcaaaagcccttcgaaaattattgtctagaaaaatgaaaactcttaggact	238	
QY	189	atccagatcccggttggagaataatgacctttcttttgagtatctctcttcaactcg	248	
Db	239	atccagatcccggttggagaataatgacctttcttttgagtatctctcttcaactcg	298	
QY	249	ttaaacacatcacaaagtttccctctatattcttccagaatcacatctcgggctct	308	
Db	299	ttaaacacatcacaaagtttccctctatattcttccagaatcacatctcgggctct	358	
QY	309	gtttgttcataattctcggagccttcctaattgcagtgaaagaaaacacagaact	368	
Db	359	gtttgttcataattctcggagccttcctaattgcagtgaaagaaaacacagaact	418	
QY	369	ctgataatattgagcgcgaataatgaactctcttagtcctcgaagacaatagcttgaact	428	
Db	419	ctgataatattgagcgcgaataatgaactctcttagtcctcgaagacaatagcttgaact	478	
QY	429	attctctcacattggtttcattcactagatcaaaaacataatttggttattctcaccaa	488	

Accession	Sequence	Position
Db	479C.....	538
Qy	469 aatagtcagtgaagctgtactgtcctgttcttgggaatttgatcacatgaact	548
Db	539:.....:.....g.....	598
Qy	549 ttcaagcattatgaattatcaattctctcgtccttccaatttgggggtgcacacag	608
Db	599:.....:.....:.....	658
Qy	609 gatttgatttgaacaagtgtttactagactgtgagaataaagatgttaataa	668
Db	659:.....:.....C.....	718
Qy	669 aaaaaaa 676	
Db	719726	

[illegible]

[illegible]

Query Match	Best Local Similarity	Score	Pred. No.	Length	Matches	Conservative	Mismatches	Indels	Gaps
Qy 9	tttcaattatccagacacacatcgtgattcagaacgacgcacagccggtgttctg	16.1%	108.8	DB 22	936	5	245	0	0
Db 59	tttcaattatccagacacacatcgtgattcagaacgacgcacagccggtgttctg	16.1%	108.8	DB 22	936	5	245	0	0
Qy 69	gtattccctccagaatactgcttcagatatggtccacagaacttccagccagacc	16.1%	108.8	DB 22	936	5	245	0	0
Db 119	gtattccctccagaatactgcttcagatatggtccacagaacttccagccagacc	16.1%	108.8	DB 22	936	5	245	0	0
Qy 129	ttttcaactcaagcccttgcaaaaatttgcctagagaaaatgaaaacttgagact	16.1%	108.8	DB 22	936	5	245	0	0
Db 179	ttttcaactcaagcccttgcaaaaatttgcctagagaaaatgaaaacttgagact	16.1%	108.8	DB 22	936	5	245	0	0
Qy 189	atccagacccgttggatgatgaccttctcttggatgatcttccacttg	16.1%	108.8	DB 22	936	5	245	0	0
Db 239	atccagacccgttggatgatgaccttctcttggatgatcttccacttg	16.1%	108.8	DB 22	936	5	245	0	0
Qy 249	ttaaacacatcccaaggtccacttatattcttcagatatccatctcgggtct	16.1%	108.8	DB 22	936	5	245	0	0
Db 299	ttaaacacatcccaaggtccacttatattcttcagatatccatctcgggtct	16.1%	108.8	DB 22	936	5	245	0	0
Qy 309	gtttgttcattaatctcgagcccttccatctgcaagaaaacacacagaact	16.1%	108.8	DB 22	936	5	245	0	0
Db 359	gtttgttcattaatctcgagcccttccatctgcaagaaaacacacagaact	16.1%	108.8	DB 22	936	5	245	0	0
Qy 369	ctgataatattgagccgagataatccttcttggtccctgagagcaatctgaaac	16.1%	108.8	DB 22	936	5	245	0	0
Db 419	ctgataatattgagccgagataatccttcttggtccctgagagcaatctgaaac	16.1%	108.8	DB 22	936	5	245	0	0
Qy 429	attctccacacatttggttcatcctagatacacaactcatttggttattccacaa	16.1%	108.8	DB 22	936	5	245	0	0
Db 479	attctccacacatttggttcatcctagatacacaactcatttggttattccacaa	16.1%	108.8	DB 22	936	5	245	0	0
Qy 489	aatagtcagtgtaagcgttactgtctgtcttgggaatttgatcaatgatgact	16.1%	108.8	DB 22	936	5	245	0	0
Db 539	aatagtcagtgtaagcgttactgtctgtcttgggaatttgatcaatgatgact	16.1%	108.8	DB 22	936	5	245	0	0
Qy 549	ttcagcattatgattatcattctcctgcttccatatttgggtgacctagag	16.1%	108.8	DB 22	936	5	245	0	0
Db 599	ttcagcattatgattatcattctcctgcttccatatttgggtgacctagag	16.1%	108.8	DB 22	936	5	245	0	0
Qy 609	gattgtgatttgtaacaatggttgtagcagcactgtagagataaagtgttcaata	16.1%	108.8	DB 22	936	5	245	0	0
Db 659	gattgtgatttgtaacaatggttgtagcagcactgtagagataaagtgttcaata	16.1%	108.8	DB 22	936	5	245	0	0
Qy 669	aaaaaaa 676	16.1%	108.8	DB 22	936	5	245	0	0
Db 719	aaaaaaa 676	16.1%	108.8	DB 22	936	5	245	0	0

ID	AAFS8254/c	AAFS8254 standard; DNA; 936 BP.
XX	AAFS8254;	
AC	AAFS8254;	
XX		
DT	24-APR-2001	(first entry)
XX		
DE	Oligonucleotide D1875.	
XX		
KM	Electron-transfer group; ETM; mismatch; genotyping;	
KW	gene expression; ss.	
XX		
OS	Synthetic.	
XX		
PN	MO200107665-A2.	
PD		
XX	01-FEB-2001.	
XX		
PF	26-JUL-2000; 2000MO-US20476.	
XX		
XX	26-JUL-1999; 99US-0145695.	
XX	17-MAR-2000; 2000US-0190259.	
XX		
PA	(CLIN-) CLINICAL MICRO SENSORS INC.	
P1		
XX	Umek RM;	
DR	WPI; 2001-159728/16.	
XX		
PT	Nucleic acids containing electron-transfer group, useful as labels in	
PT	hybridization assays, e.g. for genotyping, allowing repeat analyses on	
XX	a single surface	
XX		
PS	Example 6; Page 127; 159pp; English.	
XX		
CC	The present invention relates to a composition comprising two nucleic	
CC	acids each containing an electron-transfer group (ETM) having	
CC	different redox potentials. The invention is used for electronic	
CC	detection of nucleic acids, especially of substitutions (mismatches)	
CC	and single-nucleotide polymorphisms, e.g. for genotyping,	
XX	monitoring gene expression.	
XX		
SO	Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;	
Query Match 16.1%; Score 108.8; DB 22; Length 936;		
Best Local Similarity 0.7%; Pred. No. 5,1e-20;		
Matches 5; Conservative 418; Mismatches 245; Indels 0; Gaps		
9	ttcaaatatcacgcagaccatcatgatgttcaagcacgcacacagtcggtttctcg	68
DB	ttcaaatatcacgcagaccatcatgatgttcaagcacgcacacagtcggtttctcg	68
729	ttcaaatatcacgcagaccatcatgatgttcaagcacgcacacagtcggtttctcg	670
OY	gtatttccttcacagaatcactgcttcagaatatgagtcacagaaacttcagcacagacc	128
DB	gtatttccttcacagaatcactgcttcagaatatgagtcacagaaacttcagcacagacc	128
669	tttcaaacacaagcccttcgcaaaaattatgtctagaaaaatgaaaactttaggact	188
OY	tttcaaacacaagcccttcgcaaaaattatgtctagaaaaatgaaaactttaggact	188
DB	tttcaaacacaagcccttcgcaaaaattatgtctagaaaaatgaaaactttaggact	188
609	tttcaaacacaagcccttcgcaaaaattatgtctagaaaaatgaaaactttaggact	188
OY	tttcaaacacaagcccttcgcaaaaattatgtctagaaaaatgaaaactttaggact	188
DB	tttcaaacacaagcccttcgcaaaaattatgtctagaaaaatgaaaactttaggact	188
189	atcagaatcctggttggaaatgatgaccttcttccttgagagatccttccttcacctg	248
OY	atcagaatcctggttggaaatgatgaccttcttccttgagagatccttccttcacctg	248
DB	atcagaatcctggttggaaatgatgaccttcttccttgagagatccttccttcacctg	248
549	tttaaacacatccaaggttcccttataatcttccttcgagatccttcgagacct	308
OY	tttaaacacatccaaggttcccttataatcttccttcgagatccttcgagacct	308
DB	tttaaacacatccaaggttcccttataatcttccttcgagatccttcgagacct	308
489	gttttgctcaatattcttgagaccttcctcaatgcagctgaaaaaaccacagaact	368
OY	gttttgctcaatattcttgagaccttcctcaatgcagctgaaaaaaccacagaact	368
DB	gttttgctcaatattcttgagaccttcctcaatgcagctgaaaaaaccacagaact	368
429	tttcaaacacaagcccttcgcaaaaattatgtctagaaaaatgaaaactttaggact	188

[illegible]

[illegible][illegible]

Fri Nov 30 11:02:22 2001

us-09-735-712-9.rng

Page 12

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 00:20:23 ; Search time 83.78 Seconds
(without alignments)
1827.395 Million cell updates/sec

Title: US-09-735-712-9
Perfect score: 676
Sequence: 1 tcactctcttcaattatc.....tgyttaaataaaaaaa 676

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6C.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86.4	12.8	1669	2	US-08-916-902A-2
2	86.4	12.8	1669	2	US-09-213-389-2
3	63.2	9.3	1661	1	US-08-318-492-3
4	63.2	9.3	1661	1	US-08-707-340-3
5	63.2	9.3	1661	2	US-08-994-578-3
6	49.2	7.3	2545	1	US-07-869-933-22
7	49.2	7.3	2545	4	US-09-103-663-22
8	42.6	6.3	7218	1	US-08-232-463-14
9	34.6	5.1	246240	2	US-08-724-394A-20
10	34.6	5.1	246240	2	US-08-724-394A-21
11	34.6	5.1	246240	1	US-08-724-394A-22
12	33.6	5.0	1711	2	US-08-568-147B-1
13	33.4	4.9	6669	3	US-09-212-971-5
14	33.4	4.9	6669	3	US-08-800-929A-5
15	33.4	4.9	6669	4	US-09-617-053A-5
16	32.2	4.8	3947	4	US-08-975-762-47
17	32.2	4.8	3947	4	US-08-295-028-47
18	32.2	4.8	3947	4	US-09-106-582-47
19	32	4.7	8920	2	US-08-446-855A-1
20	32	4.7	8920	2	US-09-150-741-1
21	32	4.7	9793	1	US-08-470-202-56
22	32	4.7	9793	1	US-08-471-770-56
23	32	4.7	9793	2	US-08-468-059-56
24	32	4.7	9793	4	US-09-109-916-56
25	31.8	4.7	1400	4	US-09-041-718-1
26	31.6	4.7	4171	1	US-08-308-881-5
27	31.6	4.7	4171	2	US-09-058-263-5

28	31.6	4.7	4171	2	US-09-059-099-5	Sequence 5, Appl
29	31.6	4.7	4171	5	US-09-058-264-5	Sequence 5, Appl
30	31.6	4.7	4171	5	PCT-US95-06530-5	Sequence 5, Appl
31	31.4	4.6	3393	4	US-09-104-324B-1	Sequence 1, Appl
32	31.4	4.6	3393	4	US-09-162-713-1	Sequence 1, Appl
33	31.4	4.6	3560	4	US-09-081-345-1	Sequence 1, Appl
34	31.4	4.6	4291	2	US-08-417-210A-81	Sequence 81, Appl
35	31.4	4.6	4517	4	US-09-140-804-9	Sequence 9, Appl
36	31	4.6	785	4	US-09-276-141-11	Sequence 11, Appl
37	31	4.6	1902	4	US-08-793-044-4	Sequence 4, Appl
38	31	4.6	2839	4	US-09-061-702-1	Sequence 4, Appl
39	31	4.6	2945	3	US-09-058-489-85	Sequence 85, Appl
40	31	4.6	9439	3	US-09-058-489-89	Sequence 89, Appl
41	30.8	4.6	699	5	PCT-US95-13335-2	Sequence 2, Appl
42	30.8	4.6	2662	3	US-08-750-357-8	Sequence 8, Appl
43	30.8	4.6	3198	4	US-08-842-306B-48	Sequence 48, Appl
44	30.8	4.6	3198	4	US-08-838-973B-48	Sequence 48, Appl
45	30.8	4.6	3466	1	US-08-468-036-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-08-916-902A-2
Sequence 2, Application US/08916902A
Patent No. 5871930
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/916,902A
FILING DATE: Herewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINNOT04
CLONE: 927955
US-08-916-902A-2

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,492
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BIH94-03
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 97..741
US-08-318-492-3

Query Match	9.3%	Score 63.2	DB 1	Length 1661
Best Local Similarity	50.9%	Pred. No. 1.7e-08		
Matches 239	Conservative 0	Mismatches 213	Indels 18	Gaps 3

QY	167	aaaatgtgaatcttaggaactccaacccgtgttgaaatlaagaactttctcttg	226
Db	234	AAAATTACAAGTCTGGGGCCATCCAGATCCTGAAATGAGCAATGATCTGGCTTGGG	293
QY	227	agt-----tatcttcctttcaacctgtgtaaacacatalccaaggttcctt	277
Db	294	TGTCCTTCTGGGTTCCCTTGCAATACCCTCACTCCAAAAGCACTTTTCTTCTAC	353
QY	278	attcttcacgaatataccatctcctgggcctcgtttgttcattaaattcggacccctct	337
Db	354	CTTCTACACAGGCTACCCGATTTTGGGGTGTCTGTGTTTCTTAGTCAGGAACCTTGTTC	413
QY	338	aattgcagtgaanaaacaacagaacctgtataatltgagccgaataatgaatc	397
Db	414	TGTTGTACAGAGGATAAACCACAAGAAACATGATACAGAACAGTTTGGAAATGAACAT	473
QY	398	tcttagtgcccttagagcaatagctcggaaatactctctcaacatttggtttcaatctaga	457
Db	474	TGCCAGTCTCAAAATTTGACATAGTGGGGCTGCTTTTCTCTCAATAAATATGCACTTAA	533
QY	458	tcaaacacacatttgygtat-ctcaacaaatagtcagtgtaagctgtact----	512
Db	534	TATCCAGTCATTTAAGGAATTGTCACTCTTCATACAGAGTACCCGGACTTATGCAATTACT	593
QY	513	---glcctgtctctggaaatttgatlaacattgatacttgaaacttcaagcatatlgaaatt	568
Db	594	GGGCTCCATATCAATATGSCATGTCTCTCTACTGCTGATTTCTACACCTTGCTGAATTATG	653
QY	569	catcttcctgccttctcaatttgggggtgcaccctcagagatttgat	618
Db	654	CGTAACTATCTTACCATATAGCCATTTGGTGGCAATGCAAACTGCTGTAATT	703

```

RESULT      4
US-08-707-340-3
: Sequence 3, Application US/08707340
: Patent No. 5705615
:
: GENERAL INFORMATION:
:
: APPLICANT: Lim, Bing
:
: APPLICANT: Adra, Chaker N.
:
: TITLE OF INVENTION: Lellias, Jean-Michel
:
: TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
: ASSAYS
:
: NUMBER OF SEQUENCES: 4

```

```

CORRESPONDENCE ADDRESS:
ADDRESSSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Milltia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/707,340
FILING DATE: 03-SEP-1996
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,492
FILING DATE: 06-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/675,648
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: B1H94-03A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 97..738
US-08-707-340-3

```

Query Match	9.3%;	Score 63.2;	DB 1;	Length 1661;
Best Local Similarity	50.9%;	Pred. No. 1.7e+08;		
Matches 239;	Conservative 0;	Mismatches 213;	Indels 18;	Gaps 3
QY 167	aaaaatgaanaatcttagggactatccagatccctgtgttggaaatatagaacttctcttgg			
Db 234	AAAATTACAAGTCTTTGGGGCCATCCAGATCCTGAATGCAGCAATGATCTGGCTTGGG			
QY 227	agt-----tatcttcctttacacttgtttaaacatatccaaggttccctttat			
Db 294	TGCTCTTCTGGGTTCCCTTGCAATPACCATPACCATPCCAAAAGCACTCTTTTCTWTCAC			
QY 278	attcttccagatatccatctctgggacctgtgttggatcataaattctggagacctccat			
Db 354	CTTCTACACAGGCTACCCGATTTGGGGTGCTGTGTGTTTCTGTAGTACGGAACCTTGTC			
QY 338	aattgcagtgaaanaaaaaaacacagaaactctgataatattgaagccgataatgaact			
Db 414	TGTTGTACACAGGAGTAAACCCACCAAGAACATGAGTACGAACACTTTTGGATGAACAT			
QY 398	tcttagtgcccttgagagaatagcctggagataatcttcctccaaacttggtttaactctaga			
Db 474	TGCCAGTCTACAAATTCGACTAGTGGGGCTGCTTTTCTCTACTAATAATATGCACTTAA			
QY 458	tcaaaactacacatttggttat-tctcaaccaaaatagtcgtgaagcgttact----			
Db 534	TATCCAGTCATTAAAGGAATTTGTCACTCTTCATCAACAGTACCGGACCTATGCAATTAACAT			
QY 513	-----gtcctgtctctgggaaatttgattatacatatgatgatcttcagacatattgaaattat			
Db 594	GGGCTCCATATCAATGAGCATGGTGTCTCTACTGCTGATCTTCACCTGGCTGGAAATTAG			

QY 569 cattctcctgccttctcaatttggtgccacccaggaattgatt 618
| | | | | | | | | | | |
Db 654 CGTAACFATCTTACCATAGCCATGTGGCAATGCAAACTGCTGTAATT 703

RESULT 5
US-08-994-

Sequence 3, Application US/08994578
Patent No. 5972688
GENERAL INFORMATION:
APPLICANT: Lim, Bing
APPLICANT: Adra, Chaker N.
APPLICANT: Leilas, Jean-Michel
TITLE OF INVENTION: HTM4 METHODS OF TREATMENT AND ASSAYS
TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
Zip: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/994,578
FILING DATE: December 19, 1997
CLASSIFICATION: 433
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/707,340
FILING DATE: 03-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/675,648
FILING DATE: 03-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,492
FILING DATE: 06-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: Bih94-03a22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-9240
TELEFAX: (781) 861-9540
SEQUENCE CHARACTERISTICS:
INFORMATION FOR SEQ ID NO: 3:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 97..738
IS-08-994-578-3

	Query Match	9.3%	Score 63.2	DB 2	Length 1661
	Best Local Similarity	50.9%	Pred. No. 1.7e-08		
	Matches 239	Conservative	0	Mismatches 213	Indels 18
				Gaps	3
0Y	167	aaaatgaaaactctagagactatccagatctcgtttgaaatagacttctcttgg	226		
Db	234	AAAATTACAAGTTCTTGCGGCGCATCCAGATCTCGAATCAGCAAGATTCGTGGG	293		
0Y	227	agt-----tatctccctttacactglttaaacacatccaaaggttcccttatt	277		
Db	294	TGCTTTTCTGGGCTTCCTGCATATCCCAATACCACTTCCAAAGCAGCTTCCTTTCTTCAC	353		

QY	278	atctcttcagagatccatctctgggtcctctgttttcacatctctggacctct	337
Db	354	CTTCTACACAGGCGTACCCGGATTTTGGGGCGCTGCTTTTCTGTAGTTCAGGAACCTTGTC	4133
QY	338	aatgcagtgtaaaaaaacaacacagaaactctgtatataattgagccgaataatgaactc	397
Db	414	TGTTGTATGACAGGAGTAAACCCCAAAACATGTGATACAGAACGTTTGGAAATGAACAT	4737
QY	398	tcttagtggccctggagacaatagctctgtaataatctctccacatttggtttcactaga	457
Db	474	TGCCAGTGCTCAATTTCACACTAGTGGGGACGCTTTTCTCCTCACTAAATATAGACGTTAA	5333
QY	458	tcaaaactaacatttgggttat-ctcaccaaaatagttaagttaagcctgttaact----	5121
Db	534	TATCCAGTCATTAAAGAGATTGTACCTTTATCTACGAGATCACCCGGACCTATGCATTTACAT	5939
QY	513	---gtcctgtctcttgggaatttgcatttaacatltgatcttaacttcaagcatltaaatat	568
Db	594	GGGCTCATATCAAAATGAGCATGGTGTCTTACTGCTGATTCTCAGCTTGGTGGAATTATG	6533
QY	569	cattctctgcctcttcccaatttgggttgctgaactcagagaatttgatgt	618
Db	654	CGTACTATCTCTACCACTAGCCATGTGTGTGATGCAACAACTGCTGTAAAT	703

RESULT 6
MS-07-869-933-22

Sequence 22: Application US/07869933
Patent No. 5770396
GENERAL INFORMATION:
APPLICANT: KINET, Jean-Pierre
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
TITLE OF INVENTION: IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/869,933
FILING DATE: 19920416
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/154 NIND
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 2545 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 46..786
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 46..54
FEATURE:

NAME/KEY: mat.peptide
LOCATION: 55..786
US-07-869-933-22

Query Match 7.3%; Score 49.2; DB 1; Length 2545;
Best Local Similarity 52.3%; Pred. No. 0.00015;
Matches 158; Conservative 0; Mismatches 138; Indels 6; Gaps 2;

QY 133 caactcaagcccttcgcaaaatattctgtcagaanaatgaaatccttaggactatcc 192
DB 188 CACACAGACAGACATGCGTCATTTTGAAGAAAGAGTGGAGTTCCTGGCGTAACCC 247
QY 193 agatcgttttggaatgaactttcttttgagttatcttcccttcacactt----- 247
DB 248 AAGTTCTGTTGGTTGAATGCTTTGTTTGAAACAGTTGTCCTCCACACTCCAGA 307
QY 248 gttcaaacatataccaaggtttcccttta-tattcttcagagataccaatcttgaggt 306
DB 308 CTTCAGACTTTTGACGAGAGAGTCTTTTATATATAGACAGAGCTACCATTCCTGGGGTG 367
QY 307 cgttttgcattatcttcgtgagccttccctaattgcagtgaaagaaacacagaa 366
DB 368 CAGTCTGTTGTTGTTGCTGATTTTGTCAATATGTCGAAAGAAACACACACTGT 427
QY 367 cctcgaataatgagcgaataatgactcttccttagtgccctgagagcaatagctgaa 426
DB 428 ATCTGTTGAGAGGACCTGGAGCAACATTGTCTAGACACTGCTGCGAGCTTG6GGG 487
QY 427 tc 428
DB 488 TC 489

RESULT 7
US-09-103-663-22
Sequence 22, Application US/09103663D
Patent No. 6171803

GENERAL INFORMATION:
APPLICANT: Kinet et al.
TITLE OF INVENTION: Isolation, characterization, and use of the human beta
TITLE OF INVENTION: subunit of the high affinity receptor for
FILE REFERENCE: 50490
CURRENT APPLICATION NUMBER: US/09/103,663D
EARLIER FILING DATE: 1998-06-23
EARLIER APPLICATION NUMBER: 07/869,933
EARLIER FILING DATE: 1992-04-16
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 22
LENGTH: 2545
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: sig.peptide
LOCATION: (46)..(54)
FEATURE:
NAME/KEY: CDS
LOCATION: (46)..(786)
US-09-103-663-22

Query Match 7.3%; Score 49.2; DB 4; Length 2545;
Best Local Similarity 52.3%; Pred. No. 0.00015;
Matches 158; Conservative 0; Mismatches 138; Indels 6; Gaps 2;

QY 133 caactcaagcccttcgcaaaatattctgtcagaanaatgaaatccttaggactatcc 192
DB 188 caccacagcagacatgagcgtcatcttttgaagaagagttgaggtccgggagtaacc 247
QY 193 agatcgttttggaatgaactttcttttgagttatcttcccttcacactt----- 247

DB 248 aagttcgtttgtttgatagccttgttttggaaacagttgctgtccacactccaga 307
QY 248 gttcaaacatataccaaggtttcccttta-tattcttcagagataccaatcttgaggt 306
DB 308 ctccagactttgacgacagagtgcttttattatagacagactacatcttcgttggtg 367
QY 307 cgttttgcattatcttcgtgagccttccctaatttcagtgaaagaaacacacactg 366
DB 368 cagtcgtttgttttgcgtgacttttgcactatgtccgaaagaaacacacactg 427
QY 367 cctcgaataatgagcgaataatgactcttccttagtgccctgagagcaatagctgaa 426
DB 428 atctgttgagagcagcctgggagcaaatgttcaagcagcatcgtcgaagtttgagg 487
QY 427 tc 428
DB 488 TC 489

RESULT 8
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367

GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA

COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F15
US-08-232-463-14

Query Match 6.3%; Score 42.6; DB 1; Length 7218;
Best Local Similarity 6.6%; Pred. No. 0.015;
Matches 27; Conservative 204; Mismatches 178; Indels 0; Gaps 0;

```

Qy 1 tcatctccttcaattatccaccgacacatcagatcgaagccgacacagctccg 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1094 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1153
Qy 61 tcttctcgtatctccctccgaatcactgcttcgaaatagcccaagaacttcag 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1154 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1213
Qy 121 ccagcacttccaactcaaaagcccttgcaaaatattctcgaagaatgaatcct 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1214 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1273
Qy 181 taggacatccagactcctgcttggaatagaccttcttccttgagatcctctt 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1274 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1333
Qy 241 tcaacttgtaaacacatccaaagcttcctctatattcttcagatatcctct 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1334 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1393
Qy 301 ggggctctgttctgctcaatctcgagcctcctaattcgagtgaaagaaaccca 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1394 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1453
Qy 361 cagaacctcgataatattgagccgaataatgactctctctagtcct 409
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1454 CTTTAACTTGATGATAGTAAATTAACAGTGAATGCTTACATCCCG 1502

```

```

RESULT 9
US-08-724-394A-20
: Sequence 20, Application US/08724394A
: Patent No. 5872237
: GENERAL INFORMATION:
: APPLICANT: Feder, John N.
: APPLICANT: Krommal, Gregory S.
: APPLICANT: Lauer, Peter M.
: APPLICANT: Ruddy, David A.
: APPLICANT: Thomas, Winston
: APPLICANT: Tsuchihashi, Zenta
: APPLICANT: Wolff, Roger K.
: TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
: TITLE OF INVENTION: Sequences and Antibodies Thereeto
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
: STREET: Two Embarcadero Center, 8th Floor
: CITY: San Francisco
: STATE: CA
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/724,394A
: FILING DATE: 01-OCT-1996
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Fitts, Renee A.
: REGISTRATION NUMBER: 35,136
: REFERENCE/DOCKET NUMBER: 017957-000100
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-576-0200
: TELEFAX: 415-576-0300
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 246240 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: not relevant
: TOPOLOGY: not relevant

```

```

: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: 1..246240
: OTHER INFORMATION: /note="HLA-H.CONTIG"
US-08-724-394A-20

```

```

Query Match 5.1%; Score 34.6; DB 2; Length 246240;
Best Local Similarity 55.4%; Pred. No. 9.2;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

```

```

Qy 481 ctccacaaatagtcagtgtaagctgtactgtctcttcttggaatttgatcac 540
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36690 CTCCTCAGATCTCTGTAATTAATCTATTTATGCTGTGTAATTAACCAACTG 36749
Qy 541 tgaacttcagacttgaattatcattctccttccttcgaatttgagtgcc 600
   |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36750 TGATGCTTGAACAACAACAATAATTATTATTATTAACAATTCAAGCTCAAGTCC 36809
Qy 601 a 601
Db 36810 A 36810

```

```

RESULT 10
US-08-724-394A-21
: Sequence 21, Application US/08724394A
: Patent No. 5872237
: GENERAL INFORMATION:
: APPLICANT: Feder, John N.
: APPLICANT: Krommal, Gregory S.
: APPLICANT: Lauer, Peter M.
: APPLICANT: Ruddy, David A.
: APPLICANT: Thomas, Winston
: APPLICANT: Tsuchihashi, Zenta
: APPLICANT: Wolff, Roger K.
: TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
: TITLE OF INVENTION: Sequences and Antibodies Thereeto
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
: STREET: Two Embarcadero Center, 8th Floor
: CITY: San Francisco
: STATE: CA
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/724,394A
: FILING DATE: 01-OCT-1996
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Fitts, Renee A.
: REGISTRATION NUMBER: 35,136
: REFERENCE/DOCKET NUMBER: 017957-000100
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-576-0200
: TELEFAX: 415-576-0300
: INFORMATION FOR SEQ ID NO: 21:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 246240 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: not relevant
: TOPOLOGY: not relevant
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: 1..246240

```

Query Match	5.0%;	Score 33.6;	DB 1;	length 1711;
Best Local Similarity	47.6%;	Pred. No. 2.7;		
Matches	99;	Conservative	0;	Mismatches 109;
			Indels	0;
			Gaps	0;


```

; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF IAPs AND
; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; FILE REFERENCE: 07891/009003
; CURRENT APPLICATION NUMBER: US/09/617,053A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/800,929
; PRIOR FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 6669
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (3677)...(3951)
; OTHER INFORMATION: n can be any nucleotide
US-09-617-053A-5

```

```

Query Match 4.9%; Score 33.4; DB 4; Length 6669;
Best Local Similarity 53.4%; Pred. No. 5.2;
Matches 70; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 160 ttgctagaaaaatgaataatcttagggactatccagatccctgttggattatgaccttt 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2946 TGCGCTGTAARATGAACCTTCAAGATTTTAATTTTGTGCCCTCAAGAACTAAC 2887

QY 220 ctttggagttatcttcccttccactgtttaaaccatataccaaggttcccttataat 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2886 ATGTAGAGATTATCAACTATCACCTCCGAAGAACTTTTGACTTTGTTTCTAA 2827

QY 280 ttcttcaaga 290
    | ||| | |
Db 2826 ATGTTCTGAA 2816

```

Search completed: November 29, 2001, 00:23:11
 Job time: 9823 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 00:18:42 ; Search time 1970.99 Seconds

(without alignments)
3685.532 Million cell updates/sec

Title: US-09-735-712-9

Perfect score: 676

Sequence: 1.tcatcccttcaatratc.....tgtgttaataataaaaaaa 676

Scoring table: IDENTITY_NUC

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Listing first 45 summaries

EST:
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hrc:*
10: gb_est1:*
11: gb_est2:*
12: gb_hrc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inu:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	503.2	74.4	508	10	A1149899 g4f3h06.x
C 2	482.6	71.4	516	10	AA436088 zu03a08.r
C 3	379	56.1	387	10	AA416972 zt94h05.s
C 4	375.6	55.6	382	10	AA435988 zu03a08.s
C 5	325.2	48.1	394	10	AA758635 ah67b04.s
C 6	292.4	43.3	484	10	BE107659 UI-R-Brl-
C 7	274.8	40.7	415	10	AA470059 zt94h05.r
C 8	274	40.5	482	10	AT002083 ct38h02.s
C 9	267.4	39.6	410	10	AA411806 zt57a03.s
C 10	245.4	36.3	389	10	AA781801 ai58g01.s
C 11	148.6	22.0	538	13	AQ108532 CIT-HSP-2
C 12	148	21.9	281	10	AA707529 ab41a12.s

C 13	141	20.9	332	11	BF319786	BF319786 uy63h11.x
C 14	131.8	19.5	601	10	BE638317	BE638317 EST00003
C 15	126.2	18.7	224	10	BE638325	BE638325 EST00022
C 16	86.4	12.8	1071	10	AL544561	AL544561 AL544561
C 17	84	12.4	793	11	BG484817	BG484817 602505511
C 18	83	12.3	218	10	BE638321	BE638321 EST00013
C 19	82.2	12.2	704	11	BE638321	BE638321 602592934
C 20	80.8	12.0	517	11	BE664654	BE664654 DRABG07
C 21	80.4	11.9	508	10	AA234138	AA234138 zt51b05.r
C 22	80.4	11.9	562	10	AA418443	AA418443 zvt2e05.r
C 23	76.2	11.3	857	11	BE720182	BE720182 602692226
C 24	70.6	10.4	421	13	BE6842	BE6842 RPI011-26F1
C 25	63.6	9.4	689	11	BE719295	BE719295 602690369
C 26	63.4	9.4	899	11	BE720623	BE720623 602691889
C 27	61.2	9.1	887	10	AL531049	AL531049 AL531049
C 28	59.4	8.8	715	11	BE538851	BE538851 602568296
C 29	55.6	8.2	579	10	BE513276	BE513276 601315340
C 30	55.6	8.2	947	12	AK008652	AK008652 Mus muscu
C 31	55.4	8.2	470	10	AI807884	AI807884 w43q11.x
C 32	55	8.1	528	10	BE675149	BE675149 7f03d02.x
C 33	55	8.1	904	11	BE056859	BE056859 7k10g10.x
C 34	54.6	8.1	697	11	BE611537	BE611537 602613441
C 35	54.6	8.1	766	11	BE611537	BE611537 602613441
C 36	54.4	8.0	1504	12	AK003110	AK003110 Mus muscu
C 37	53	7.8	1184	12	AK017928	AK017928 Mus muscu
C 38	52.6	7.8	515	11	BE434337	BE434337 602506331
C 39	52.6	7.8	594	10	AV717594	AV717594 AV717594
C 40	52.6	7.8	597	10	AV716310	AV716310 AV716310
C 41	52.6	7.8	619	10	AV661743	AV661743 AV661743
C 42	52.6	7.8	680	11	BE546745	BE546745 602574115
C 43	52.6	7.8	735	10	AV715678	AV715678 AV715678
C 44	52.4	7.8	504	11	BE522968	BE522968 UI-R-C2P-
C 45	52.4	7.8	755	10	AI950360	AI950360 wpl0b06.x

ALIGNMENTS

RESULT 1
A1149899/c
LOCUS g4f3h06.xl 508 bp mRNA EST 10-NOV-1998
DEFINITION g4f3h06.xl Soares testis NHT Homo sapiens cDNA clone IMAGE:1752827
3' similar to: SM:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;
mRNA sequence.

ACCESSION A1149899
VERSION A1149899.1 GI:3678368
KEYWORDS EST.

ORGANISM human:
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 508)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@remail.nih.gov

COMMENT CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldó
Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNI at:

www.bio.litl.gov/dbtp/Image/Image.html
Insert length: 742 Std Error: 0.00

Seq primer: -40ml3 fwd. EF from Amersham
High quality sequence stop: 455.

FEATURES Location/Qualifiers
1..508

Source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1752827"

```

/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
Inc., and primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGCGCCGCCAATTTTATGACCTTTCTTTT
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      186 a      91 c      89 g      142 t
ORIGIN

```

```

Query Match      74.4%; Score 503.2; DB 10; Length 508;
Best Local Similarity 99.4%; Pred. No. 4.3e-102;
Matches 505; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

165 agaaaaaagaaactttaggagactacagatccctgtttggaattgaccttctt 224
|||||
508 AGAAAAAGAAATCTTAGGACATCCATGTTGGAATATGACCTTTCTTTT 449
|||||
225 ggaattatcttcctttacacttgtaaacacatccaaagttcccttattctt 284
|||||
448 GGAGTTATCTTCCTTTTACCTTGTAAACCAATACCAAGGTTCCCTTATATTTCTT 389
|||||
285 tcagatatcatcttgaggctctgtttgttcaataattcggagccttcaattgca 344
|||||
388 TCAGATATTCATCTTGCGGCTCTGTTTGTTCATTAATTCGAGCCTTCCTTAATTGCA 329
|||||
345 gtgaaaagaaacacacagaaacttgataatattgagccgaataatgaattctt 404
|||||
328 GTGAAAAAGAAACACAGAAACCTGATTAATATGAGCCGAAATGAAATTTCTTGT 269
|||||
405 gccctgagagaataagctggaatacttctctcaattgttttcatctgataaac 464
|||||
268 GCCCTGGAGCAATAGCTGGAATCTCTCCATCATTTGGTTTCATCTAGATCAAAAC 209
|||||
465 tacatttggtattctcaccacaaatagtcagtgttaagctgtactgcctgtctt 524
|||||
208 TACATTTGTGTTATTCCTACCAAAATAGTCAGTAAAGCTTACGCTCTGTTCTTG 149
|||||
525 ggaatttgattacatgatacttcgaactatgaactatcaattccttccttcc 584
|||||
148 GGAATTTTGAATTACATTCATGACCTTCAGCATTTATTTGAAATTCATCTCTG 89
|||||
585 tcaatttgggggtccactcagaagatgtgattgtgaacaatgttggatcagcact 644
|||||
88 TCAATTTTGGGGTCCACTCAGAGATGTGATTTGTGAACAATGTTGATGACACTG 29
|||||
645 tgaataaagaatgtgttaaaaaaa 672
|||||
28 TGAGAAATTAAGATGTGTAATAATAAAA 1
|||||

RESULT 2
AA436088      516 bp      mRNA      EST      09-NOV-1997
LOCUS      2U03a08.r1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:730742
DEFINITION      5'. similar to SM:CD20_HUMAN P11836 B-LYMPHOCTYE ANTIGEN CD20 ;'
mRNA sequence.
ACCESSION      AA436088
VERSION      AA436088.1 GI:2141002
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Joat, S.,
AUTHORS

```

```

TITLE
JOURNAL
COMMENT
Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 733 Std Error: 0.00
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 482.

FEATURES
Source
1..516
/organism="Homo sapiens"
/db_xref="GDB:5927437"
/db_xref="taxon:9606"
/clone="IMAGE:730742"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
Inc., and primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGCGCCGCCAATTTTATGACCTTTCTTTT
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      149 a      117 c      80 g      170 t
ORIGIN

Query Match      71.4%; Score 482.6; DB 10; Length 516;
Best Local Similarity 99.2%; Pred. No. 1.6e-97;
Matches 485; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 tcaattccttcaaatattacacgacacatcagatgaatcaagcagcacagctcgg 60
|||||
20 TCATCTCTCTTCAATTAATATACCGACACATCAGATGATTCAAGCACCCACACATCCGG 79
|||||
61 tgtttcgtatttccctcagaataactgcttcagaatagatgagtcacagaacttcag 120
|||||
80 TGTTTCTGTATTTCTCTCCAGAAATCACTGCTTCAGATTAATGAGTCCACAGAACTTTAG 139
|||||
121 ccacgaccttccaactcaaaagcccttgcacaaatatttgcctagaacaaatgaatct 180
|||||
140 CCACGACCTTTTCAACTCAAAAGCCCTTGCAAAATTTTGTGTAGAAATTAATTAATCT 199
|||||
181 taggacataccagatcctgtttggaattatgaacttcttcttggagatattctctt 240
|||||
200 TAGGACTATCTCAGATCTCTGTTTGGAAATTAATGACCTTTCTTTTGGAGTTATCTTCT 259
|||||
241 tcacctgttaaaacacatccaaggttcccttatattcttccatcaagatcatctc 300
|||||
260 TCACCTTTTAAACCAATCCAGATTCCTTTATATTTCTTTTCAGAGATATTCATTTCT 319
|||||
301 ggggctctgtttgttcaatattcttgagccttccatattgcagtgaacaaaccca 360
|||||
320 GGGGCTCTGTTTGTCTATTAATCTGGAGCCTTCCTTAATTCAGTGAATAAAGAAACCA 379
|||||
361 cagaacacttgataatattggccgaataatgaatcttcttaagtcctgagcaatag 420
|||||
380 CAGAAACTGTATATATATGAGCGGAATTAATTAATTTCTTAGGCCCTGGAGCAATAG 439
|||||
421 ctggaatcatctctcactcattgttgcattccctcagatcaaaactaattgttgatt 480
|||||

```

```

Db      440 CTGGAAATCTCTCTACATTTGTTTCATCTCATGATCAAAACATATTGTGTTATT 499
QY      481 CTCACCAAA 489
Db      500 CTCACAAA 508

RESULT 3
AA416972/c 387 bp mRNA EST 09-NOV-1997
ZU09405.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730041
3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;
mRNA sequence.
ACCESSION AA416972
VERSION AA416972
KEYWORDS AA416972.1 GI:2077080
SOURCE EST.
ORGANISM human.
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 387)
          Hillier,L., Allen,M., Bowles,L., Dubuque,T., Gelsel,G., Jost,S.,
          Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
          J., Moore,B., Schellenberg,K., Stepien,M., Tan,F., Theising,B.,
          White,Y., Wylie,T., Waterston,R. and Wilson,R.
          WashU-NCI human EST Project
          Unpublished (1997)
          CONTACT: Wilson RK
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: estewatson.wustl.edu
          This clone is available royalty-free through LNL ; contact the
          IMAGE Consortium (info@image.llnl.gov) for further information.
          Insert Length: 656 Std Error: 0.00
          Seq primer: -41m13 fwd. ET from Amersham.
FEATURES
          source
          1..387
          /organism="Homo sapiens"
          /db_xref="GDB:5928570"
          /db_xref="taxon:9606"
          /clone="IMAGE:730041"
          /clone_1ib="Soares_testis_NHT"
          /sex="male"
          /lab_host="DH10B"
          /note="Vector: pRT73D-Pac (Pharmacia) with a modified
          polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
          was prepared from mRNA obtained from Clontech Laboratories
          , Inc., and primed with a Not I - oligo(dT) primer [5'
          TGTTCACATCTGAGTGGAGCGCGCCCAATTTTCTTTTCTT 3'].
          Double-stranded cDNA was ligated to Eco RI adaptors
          (Pharmacia), digested with Not I and cloned into the Not I
          and Eco RI sites of the modified pRT73 vector. Library
          went through one round of normalization to Cot5, and was
          constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 138 a 77 c 67 g 105 t
ORIGIN
Query Match 56.1%; Score 379; DB 10; Length 387;
Best Local Similarity 98.7%; Pred. No. 2e-74;
Matches 382; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

Db      267 TAGTCCCTGGAGCAATAGCTGGAATATTCCTCCACATTTGTTATCTTAGATCA 208
QY      461 aaactacattgtgtatctcaccacaatagtcagtcgtaagcctgtactctgtt 520
Db      207 AAACATCATTTGTGTTATTCCTCCACCAAAATAGCTAGTAAAGCGTGTACTCTGTT 148
QY      521 ctgggaatttgattacattgacttcagcttatgaaatttctctcgc 580
Db      147 CTGGGAATTTGATTCATGATGACTTTCACCATTTATGAAATTCATTTCTCTGCC 88
QY      581 ttctcaatttgggtgcacacacagagatttgatttgacaacatgtgttgaactagc 640
Db      87 TTTCTCAATTTTGGGGTGCACACAGAGATTGTGATTTGAGACAAATGTTGTGACTAG 28
QY      641 actgtgagataaagaatgtgttaaat 667
Db      27 ACTGTGAGATAAAGATGTGTTAAAT 1

RESULT 4
AA435988/c 382 bp mRNA EST 09-NOV-1997
ZU03208.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730742
3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;
mRNA sequence.
ACCESSION AA435988
VERSION AA435988
KEYWORDS AA435988.1 GI:2140902
SOURCE EST.
ORGANISM human.
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 382)
          Hillier,L., Allen,M., Bowles,L., Dubuque,T., Gelsel,G., Jost,S.,
          Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
          J., Moore,B., Schellenberg,K., Stepien,M., Tan,F., Theising,B.,
          White,Y., Wylie,T., Waterston,R. and Wilson,R.
          WashU-NCI human EST Project
          Unpublished (1997)
          CONTACT: Wilson RK
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: estewatson.wustl.edu
          This clone is available royalty-free through LNL ; contact the
          IMAGE Consortium (info@image.llnl.gov) for further information.
          Possible reversed clone: similarity on wrong strand
          Insert Length: 733 Std Error: 0.00
          Seq primer: -41m13 fwd. ET from Amersham
          High quality sequence stop: 280.
FEATURES
          source
          1..382
          /organism="Homo sapiens"
          /db_xref="GDB:5927437"
          /db_xref="taxon:9606"
          /clone="IMAGE:730742"
          /clone_1ib="Soares_testis_NHT"
          /sex="male"
          /lab_host="DH10B"
          /note="Vector: pRT73D-Pac (Pharmacia) with a modified
          polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
          was prepared from mRNA obtained from Clontech Laboratories
          , Inc., and primed with a Not I - oligo(dT) primer [5'
          TGTTCACATCTGAGTGGAGCGCGCCCAATTTTCTTTTCTT 3'].
          Double-stranded cDNA was ligated to Eco RI adaptors
          (Pharmacia), digested with Not I and cloned into the Not I
          and Eco RI sites of the modified pRT73 vector. Library
          went through one round of normalization to Cot5, and was
          constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 135 a 73 c 67 g 107 t
ORIGIN

```

Query Match 55.6%; Score 375.6; DB 10; Length 382;
 Best Local Similarity 99.0%; Pred. No. 1.1e-73;
 Matches 378; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

291 tatcatcttgaggctgctgtttgttcaatattctgagcctctcttaattgcaatgaa 350
 |||||||
 302 TATCATCTTGAGGCTGCTGTTTGTTCATTAAATCTGGAGCCCTCTAATTCAGTGA 323
 |||||||
 351 agaaaacacagaaactctgataatattgagccgaataatgactcttctagtcctg 410
 |||||||
 322 AGAAAACACAGAAACTCTGATATATGAGCCGATATGATTTTCTTAGTGCCTG 263
 |||||||
 411 agagcaataagctggaataattctctcacattgttccatctctagatcaaaactac 470
 |||||||
 262 GGAGCAATAAGCTGGAATATCTCTCCATCTTGGTTTCATCTAGATCAAAACTAC 203
 |||||||
 471 tctgttcttctcaccacaataagctgtaagctgttactgtctgttcttggaatt 530
 |||||||
 202 TCTGTATTCTCACCACAAATAGTCAAGTGAAGCTGTACTGCTCTGTTGGGAAT 143
 |||||||
 531 ttgattacattgactctcagcattatgaaatattcattctctgctcttccaatt 590
 |||||||
 142 TTGATTACATTGATGACTTTTGAGCATTAATGATTAATTCATTTCTGCTTTCAT 83
 |||||||
 591 ttgggtgctccactcagaggaattgtaattgtaacaagtgttctgactgagactgtg 650
 |||||||
 82 TTGGGCTCCACTCAGAGATTTGATTTGTGAAACATGTTGTGACATGACATGAGAA 23
 |||||||
 651 taagaatgtttaaataaaaa 672
 |||||||
 22 TAAAGATGTGTTAAATATATA 1
 |||||||

RESULT 5
 AA758635/c mRNA EST 29-DEC-1998
 LOCUS ab67b04.81 Soares testis_NHT Homo sapiens cDNA clone 1320655 3'
 DEFINITION similar to SW:CD20_HUMAN F11836 B-LYMPHOCTE ANTIGEN CD20 ;, mRNA
 sequence.
 ACCESSION AA758635.1 GI:2806498
 VERSION AA758635
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 394)
 NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-remail.nih.gov
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
 , Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
 www-bio.llnl.gov/bdrrp/image/image.html
 Insert Length: 661 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 390.
 Location/Qualifiers
 1..394
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="1320655"
 /clone_id="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc., and primed with a Not I - oligo(dT) primer [5'
 TGGTCAATCTGGAAGTGGAGCGGCCGCCCAATTTTCTTTTCTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 133 a 73 c 67 g 121 t
 ORIGIN

Query Match 48.1%; Score 325.2; DB 10; Length 394;
 Best Local Similarity 95.7%; Pred. No. 1.8e-62;
 Matches 377; Conservative 0; Mismatches 13; Indels 4; Gaps 4;

281 tcttcaggaatccattctggtggtctgtttgttcaatattctgagccttccat 340
 |||||||
 394 TCTTTCAGGATATTCATTTCTGGGCTCTGTATGTATCAATTAATCTGGAGCTTCT 335
 |||||||
 341 tgcagtgaaaagaaaacacagaaactctgataatttgagccgaataatgaactct 400
 |||||||
 334 TGCAGTGAAGAAAGAAACACAGAAACTCTGATATATGAGCCGATATGAATTTCT 275
 |||||||
 401 tagtgccctgagagaatagctggaatcattctctcacaattg-gttccatctagac 459
 |||||||
 274 TAGTGCCCTGGAGAGCAATAGCTGGAATCATCTCTACATTTGATTTGATTCCT 215
 |||||||
 460 aaactaatattgtgttattctacaaaatgaagtcagtgtgaagctgttactgtc 519
 |||||||
 214 AAACCTAATTTGTATTATTTCTCACAAATATGACAGTGAAGCTGTACTGTCTGA 155
 |||||||
 520 tcttggaatttgaatcattgactctcaagcat-tatgaaattcaat-cttctct 577
 |||||||
 154 TCTTGGGAATTTAGAAATGATGATGACTTTGAGATATATGATTTCTATTTCT 95
 |||||||
 578 gccctt-ctcaatttgggggtgcactcagaaggaattgattggaacaattgttga 636
 |||||||
 94 GCCTTTACTCAATTTTGGGGTCCACTCAGAGGATTTGATTTGATGACATGATG 35
 |||||||
 637 tagcactgtaagaataagatgtgttaataaaa 670
 |||||||
 34 TAGCACTGTGAGAAATTAAGATATGTTAAATATA 1
 |||||||

RESULT 6
 BE107659/c mRNA EST 13-JUN-2000
 LOCUS BE107659.1 GI:8499769
 DEFINITION UT-R-BT1-ame-d-04-0-01.s1 UT-R-BT1 Rattus norvegicus cDNA clone
 BE107659
 ACCESSION BE107659.1 GI:8499769
 VERSION BE107659
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 484)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu


```

QY 181 taggactatccagatcctggttgtaattacacattcttcttgaggtatcttctt 240
    |||
Db 205 TAGG----- 208
QY 241 tcacctgtttaaaccatataccaagttcccttataattcttccagatataccattct 300
    |||
Db 209 -----GGATATCCATTCT 221
QY 301 ggggctctgtttgttcaatattctgagccttccttaattgacgtggaagaagaacca 360
    |||
Db 222 GGGGCTCTGTTTGTTCATTATTCGTGAGCCTTCCTTAATTCAGAGTGAAGAAAGAACCA 281
QY 361 cagaactctgataataattgagccgaataatgaatcttcttagtgccttgagagaatag 420
    |||
Db 282 CAGAACTCTGATTAATATGTGAGCCGAATTAATGAAATTTCTTAGTCCCTGGAGACAATAG 341
QY 421 ctggaatcattctcctacattggtttctcctcagatcaactacattgtgtatt 480
    |||
Db 342 CTGGAATCATTCTCTCAGATTGTTGTTCACTGATCAAACTACATTGTTGTTATT 401
QY 481 ctacccaataagt 494
    |||
Db 402 CTCACCAAAATAGT 415

```

```

RESULT 8
LOCUS AI002083 482 bp mRNA EST 27-AUG-1998
DEFINITION ct38b02.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1619019
ACCESSION AI002083
VERSION AI002083.1 GI:3202120
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 482)
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-f@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www.bio.lnl.gov/dbp/image/image.html
Insert Length: 598 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 393.
Location/Qualifiers
1.482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1619019"
/clone_11b="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73d-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I. Site 2: Eco RI. 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I oligo(dt) primer '5'
TGTTCACCAATCTGAGTGGAGCGGCCCAATTTTGTGTGTGT 3'}.
Double-stranded cDNA was ligated to Eco RI adaptors.
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

```

FEATURES

source

```

1.482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1619019"
/clone_11b="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73d-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I. Site 2: Eco RI. 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I oligo(dt) primer '5'
TGTTCACCAATCTGAGTGGAGCGGCCCAATTTTGTGTGTGT 3'}.
Double-stranded cDNA was ligated to Eco RI adaptors.
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

```

```

BASE COUNT 157 a 85 c 99 g 141 t
ORIGIN

```

```

Query Match 40.5%; Score 274; DB 10; Length 482;
Best Local Similarity 71.9%; Pred. No. 4.3e-51;
Matches 482; Conservative 0; Mismatches 0; Indels 188; Gaps 2;

```

```

QY 1 tcattccttcaaatatccacgacacacatcagatcagacccgacacagctccg 60
    |||
Db 482 TCATCTCTCTTCAATATATATACCGACACATCATGATTCAGACCGCACAGTCCGG 423
QY 61 tttcttgatatttccctcagaaatcactgcttcaagaatagttccagaaacttcag 120
    |||
Db 422 TGTTCGTGATTTCTTCACCAAAATCACTGCTCAGATATGATGACCAAGAACTTCAG 363
QY 121 ccagaccttccacccaagcccttgcaaatatttgcagaaaaatgaaatct 180
    |||
Db 362 CCAGACCTTTTCAACTCAAGCCCTTGGCAAAATATTTGCTAGAAAAATGAATATCT 303
QY 181 taggactatccagatcctggttgtaattacacattcttcttgaggtatcttctt 240
    |||
Db 302 TAGG----- 297
QY 241 tcacctgtttaaaccatataccaagttcccttataattcttccagatataccattct 300
    |||
Db 298 -----GGATATCCATTCT 286
QY 301 ggggctctgtttgttcaatattctgagccttccttaattgacgtggaagaagaacca 360
    |||
Db 285 GGGGCTCTGTTTGTTCATTATTCGTGAGCCTTCCTTAATTCAGTGAAGAAAGAACCA 226
QY 361 cagaactctgataataattgagccgaataatgaatcttcttagtgccttgagagaatag 420
    |||
Db 225 CAGAACTCT----- 214
QY 421 ctggaatcattctcctacattggtttctcctcagatcaactacattgtgtatt 480
    |||
Db 215 -----GATCAAACTACATTGTTGTTATT 191
QY 481 ctacccaataatagtcagtgtaagcgttactgctccgttcttggaatttgatcat 540
    |||
Db 190 CTCACCAAAATAGTCAGTGTGAGCCTGTACTGCTGCTTCTTGGAATTTGATTACAT 131
QY 541 tgaatccttcagcatatgaattatcattctccttccattcaatttgggtgcc 600
    |||
Db 130 TGATGACTTTCAGCATATGATTAATTAATTCATTCTGCTTCTCAATTTGGGGTGCC 71
QY 601 actaagagatgtgattgtaagaatggttctgactagacactgtgagataagatgtg 660
    |||
Db 70 ACTCAGAGGATGTGATTTGTAACAAATGTTGTGACAGCACTGTGAATTAAGATGTG 11
QY 661 ttaaatcaaa 670
    |||
Db 10 TTAATAATAAA 1

```

```

RESULT 9
LOCUS AA411806 410 bp mRNA EST 12-AUG-1997
DEFINITION zt67a03.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:727372
ACCESSION AA411806
VERSION AA411806.1 GI:2070377
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 410)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Maitra, M., Martin, J., Moore, B.

```

TITLE
 JOURNAL
 COMMENT
 'Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, F., Meterson, R. and Wilson, R.
 Washu-Merck EST Project 1997
 Unpublished (1997)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 608 Std Error: 0.00
 Seq primer: -41m13 fwd. ET from Amersham
 High quality sequence stop: 401.

FEATURES

source

1. 410
 /organism="Homo sapiens"
 /db_xref="GDB:5924284"
 /db_xref="taxon:9606"
 /clone="IMAGE:727372"
 /clone_lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 Inc., and primed with a Not I - oligo(dT) primer [5'
 TGTTCACATCTGAGTGGAGCGGCCGCCCAATTTTCTTTTGT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
 ORIGIN
 147 a 68 c 119 t

Query Match 39.6%; Score 267.4; DB 10; Length 410;
 Best Local Similarity 96.1%; Pred. No. 1.3e-49;
 Matches 274; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 113 atttcaagcagcagcctttcaaccagaagcccttgcaaaatatttgcagaaat 172
 |||||
 DB 410 ACTTTACGACGACGACCTTTTCAACCAAGCCCTTGCAGAAATATTGCTAGAGAAAT 351
 |||||
 QY 173 gaaatcttaggactatccagatccctgttggaattatgacctttctttgagttat 232
 |||||
 DB 350 GAAATCTTAGGAGCTATCCAGATCTCTGTTGGAATTATGACCTTTCTTTGGAGTAT 291
 |||||
 QY 233 ctctctttacacttgtaaaacacatccaaagttcccttattatctttcagagata 292
 |||||
 DB 290 CTTCCTTTACCTTGTAAACCAATCCAAAGTTCCCTTATATTTCTTTCAAGGATA 231
 |||||
 QY 293 tccattctggagctctgtttgtcatatattcttgagccctcccaatgcagtaaaag 352
 |||||
 DB 230 TCCATTCTGGGCTCTGTTTCTTCAATATCTGAGGCTTCCATATGACGTAAGG 171
 |||||
 QY 353 aaaaaccagagaactctgataatattgagccgaataatgaactc 397
 |||||
 DB 170 AAAAAACACAGAACTCTGGCAATTTGTATGATGATGACTTT 126

RESULT 10
 AA781801/c 389 bp mRNA EST 31-DEC-1998
 LOCUS AA781801 a158q01.s1 Soares_testis_NHT Homo sapiens cDNA clone 1375056 3'
 DEFINITION similar to SM:CD20_HUMAN P11836 B-LVMPHOCYTE ANTIGEN CD20 ;, mRNA
 sequence.
 ACCESSION AA781801
 VERSION AA781801.1 GI:2841132
 KEYWORDS human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 389)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

JOURNAL

COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
 , Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/dbrr/image/image.html
 Insert Length: 983 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 311.

FEATURES

source

1. 389
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="1375056"
 /clone_lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 Inc., and primed with a Not I - oligo(dT) primer [5'
 TGTTCACATCTGAGTGGAGCGGCCGCCCAATTTTCTTTTGT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
 ORIGIN
 140 a 66 c 115 t

Query Match 36.3%; Score 245.4; DB 10; Length 389;
 Best Local Similarity 95.8%; Pred. No. 1e-44;
 Matches 252; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 135 actcaagcccttgcanaaattatttgcagaaaatgaaatcttaggactaccag 194
 |||||
 DB 389 ACTCAAGCCCTTGCAAAAATATTGCTAGAGAAATGAAATCTTAGGACTATCCAG 330
 |||||
 QY 195 atccgtttggaattatgacctttctttggaattatcttcttacccttgtaaaa 254
 |||||
 DB 329 ATCCGTTTGGAAATATGACCTTTCTTTGGAGTATCTTTCCTTTGACCTGTAAATA 270
 |||||
 QY 255 caatccaaagttcccttattatcttcaagatatccattctggagcctgttttg 314
 |||||
 DB 269 CCATATCCAGATTTCCCTTATATTCTTTACAGATATCCATTTCTGGGCTCTGTTTG 210
 |||||
 QY 315 ttcaatattctgagccttccaaatlgcagtgaaaagaaaacacagaaacttgata 374
 |||||
 DB 209 TTCATTAAATTCGAGGCTTCTTAATTGAGTGAAGAAACACAGAAACTCTGGGA 150
 |||||
 QY 375 atattgagccgataatgaactc 397
 |||||
 DB 149 ATTTGATTACATGATGACTTT 127

RESULT 11
 AQ108532 538 bp DNA GSS 29-AUG-1998
 LOCUS AQ108532 CIT-HSP-237967.TP CIT-HSP Homo sapiens genomic clone 237967, DNA
 DEFINITION sequence.
 ACCESSION AQ108532

VERSION	AO108532.1	GI:3485222
KEYWORDS	GSS.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 538) Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.	
TITLE	Use of a random human BAC End Sequence Database for Sequence-Ready Map Building	
JOURNAL	Unpublished (1998)	
COMMENT	other_GSSs: CIT-HSP-237967..TR Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: mdadams@tigr.org Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html. Seq primer: MJ3-21 Class: BAC ends.	

FEATURES	SOURCE	LOCATION/QUALIFIERS
		1. 538
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="237967"
		/clone_lib="CIT-HSP"
		/sex="Male"
		/cell_type="Sperm"
		/note="Vector: pBeloBAC11; site_1: HindIII; site_2
		HindIII"
BASE COUNT		140 a 108 c 115 g 175 t
ORIGIN		

Query Match	Score	148.6	DB 13	Length 538
Best Local Similarity	94.5%	Pred. No. 3,3e-23		
Matches 154	Conservative	0	Mismatches 9	Indels 0
0Y	506	tgttactgtccctgtgtctcttgggaatttgatatacatgatgacttccaattgaatt	565	
Db	152	tcttctcttctcttattacaggggaatttttgatttgcatttgatgactttcagcatttattgatt	211	
Y	566	attcaattctctgcgccttctcaatttgggtggccactcagaaggatttggatttggaaaca	625	
Db	212	attcattttctctgccttcttctcaatttttggggggccacttcagagatttggatttggaaaca	271	
0Y	626	atgttgttgaactagaacctgttggaaataaagatgtgttaaaata	668	
Db	272	atgtgttggactagacacgtgcagattaaagattgtttaaata	314	

RESULT 12	LOCUS	AA707529/c	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
	AA707529	281 bp	mRNA	EST	12-JAN-1999		human.	Homo sapiens			
	aa1a12.s1	Scarses Testis_NHT	Homo sapiens cDNA clone 1292062 3',								
	AA707529		mRNA sequence.								
	AA707529.1	GI:2717447					EST.				
							human.				
								Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
								Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
								1 (bases 1 to 281)			
								NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap .			
								National Cancer Institute, Human Genome Anatomy Project (CGAP),			
								Tumor Gene Index			

JOURNAL
 COMMENT
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgagabs-r@mail.nih.gov
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
 , Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution
 Information can be
 found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/dbtrp/image/image.html
 Insert Length: 456 Std Error: 0.00
 Seq primer: ~40ml3 fwd. ET from Aromsham
 High quality sequence stop: 218.
 Location/Qualifiers
 1. 281
 /contig=Homo_Gad1000

FEATURES	SOURCE
Location/Qualifiers	1..281
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="1292062"	
/clone_lib="Soares_testis_NHT"	
/sex="male"	
/lab_host="DH10B"	
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories , Inc." and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAACTGGAGCGCGCCCAATTTTTTTTTTTTTTTT 3']	
(Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73D vector. Library went through one round of normalization to Cots, and was constructed by Bento Soares and M. Fatima Bonaldo."	
BASE COUNT	94 a
ORIGIN	54 c 45 g 88 t

Query Match	Similarity	21.9%	Score 148	DB 10	Length 281
Best Local Match	Similarity	100.0%	Prod. NO. 5.2e-23		
Matches 148	Conservative	0	Mismatches	0	Gaps
				Indels	
Oy	523	tgggaatttattacattacattgactgacttccaacathtaaattacattctctgcctt	582		
Db	148	tgggaatttattacattacattgacttccaacathtaaattacattctctgcctt	89		
Oy	553	tctcaattttgggtggccacttcagagagatttgtatttggacaatgttttgactagcac	642		
Db	88	tctcaattttgggtggccacttcagagagatttgtatttggacaatgttttgactagcac	29		
Oy	643	tgtggaataaagatgtgtttaataaa	670		
Db	28	tgtggaataaagatgtgtttaataaa	1		
RESULT	13				
LOCUS	BF319786/c				
DEFINITION	uv63h11.x1 McCarrey Eddy round spermatid Mus musculus cDNA clone IMAGE:3664293 3', mRNA sequence.				
ACCESSION	BF319786				
VERSION	BF319786.1				
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 332)				
	Maria,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wyllie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritzer,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.				
TITLE	The WashU-NCI Mouse EST Project 1999				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Maria M/WashU-NCI Mouse EST Project 1999				

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2001, 21:32:37 ; Search time 1789.67 Seconds

(without alignments) 4148.100 Million cell updates/sec

Title: US-09-735-712-7

Perfect score: 450

Sequence: 1 atgtatcgaacgacgcgcaca.....attgtgaacaattgttga 450

Scoring table: IDENTITY_NUC

Searched: 1472140 seqs, 8248589755 residues

al number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

GenEmbl:
1: gb_ba:*
2: gb_bt:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rnd:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rnd:*
36: em_htg_other:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	450	100.0	450	6	AX179787 Sequence
2	347.4	77.2	603	6	AX179781 Sequence
3	347.4	77.2	606	6	AX179789 Sequence
4	347.4	77.2	691	9	AB013103 Sequence
5	347.4	77.2	694	9	AF237907 Sequence
6	347.4	77.2	697	9	AF321127 Sequence
7	340.8	75.7	417	6	AX179785 Sequence
8	153.4	34.1	138097	2	AC027787 Sequence
9	153.4	34.1	138097	2	AC027787 Sequence
10	153.4	34.1	147788	2	AC015840 Sequence
11	153.4	34.1	161039	2	AP003127 Sequence
12	153.4	34.1	166804	2	AP001034 Sequence
13	153.4	34.1	166820	2	AC090401 Sequence
14	132.4	29.4	161039	2	AP003127 Sequence
15	113	25.1	246	6	AX179783 Sequence
16	112.4	25.0	166804	2	AP001034 Sequence
17	69.8	15.5	2805	9	AF367473 Sequence
18	68.4	15.2	1076	9	AF068288 Sequence
19	66.8	14.8	720	9	AB022821 Sequence
20	66.8	14.8	916	9	AB013102 Sequence
21	66.8	14.8	1619	9	AF237912 Sequence
22	66.8	14.8	1669	6	AR035695 Sequence
23	66.8	14.8	1669	6	AR083580 Sequence
24	50.2	11.2	708	6	AX101306 Sequence
25	50.2	11.2	1010	10	MUSFCERB Sequence
26	46	10.2	1628	9	BC008487 Sequence
27	46	10.2	1646	9	HUMIERB Sequence
28	46	10.2	1661	6	A68627 Sequence
29	46	10.2	1661	6	AR082014 Sequence
30	46	10.2	1661	6	I25809 Sequence
31	44.8	10.0	165139	9	HS436K10 Sequence
32	44.4	9.9	2542	10	RATFCIGA Sequence
33	44.4	9.9	2545	6	AR123798 Sequence
34	44.4	9.9	2545	6	I09687 Sequence
35	43.4	9.6	518	10	AF280401 Sequence
36	43.2	9.6	180017	2	AC073277 Sequence
37	43.2	9.6	191916	2	AC068760 Sequence
38	43	9.6	12029	3	AE001379 Sequence
39	43	9.6	141540	9	AL138828 Human DNA
40	43	9.6	175238	2	AC022654 Human DNA
41	42.4	9.4	11002	5	CHKUPAG Chicken uro
42	42.4	9.4	165358	2	AC009020 Human DNA
43	42.4	9.4	186187	9	AC007464 Human DNA
44	42	9.3	61010	9	AL138961 Human DNA
45	42	9.3	167830	2	AC027768 Homo sapi

ALIGNMENTS

RESULT 1
AX179787 450 bp DNA PAT 06-AUG-2001
LOCUS Sequence 7 from Patent WO0146417.

DEFINITION AX179787
ACCESSION AX179787
VERSION AX179787.1 GI:15132147

KEYWORDS
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 450)
AUTHORS Walke,D.W. and Turner,C.A.
TITLE Human membrane proteins and polynucleotides encoding the same
JOURNAL having homology to cd20 proteins and ige receptors
Patent: WO 0146417-A 7 28-JUN-2001;
Lexicon Genetics Incorporated (US)

FEATURES
source 1..450
/organism="Homo sapiens"


```

|||||
Db 273 TTTATATTTCTTTCAGAGATATTCATCTGGGCTCTGTTTGTCATTAATCTGGAGCC 332
QY 301 ttcccaatgcagtgaaagaaacacagaaactctcgggaatttggttcacatgag 360
Db 333 TTCCTAATTTGACGTAGAAAAGAAACACAGAACTCTGATATATTTGAGCCGAATTAATG 392
QY 361 acttc 365
Db 393 AATCT 397

RESULT 4
LOCUS AB013103 691 bp mRNA PRI 20-MAR-2001
DEFINITION Homo sapiens mRNA for MS4A5, complete cds.
ACCESSION AB013103
VERSION AB013103.1 GI:11559213
KEYWORDS MS4A5; CD20-like 2.
SOURCE Homo sapiens testis cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 Ishibashi, K., Suzuki, M., Sasaki, S. and Imai, M.
AUTHORS Identification of a new multigene four-transmembrane family (MS4A)
TITLE related to CD20, Htm4 and beta subunit of the high-affinity Ige
JOURNAL receptor
MEDLINE Gene 264 (1), 87-93 (2001)
REFERENCE 2 (sites)
AUTHORS Ishibashi, K., Sasaki, S. and Marumo, F.
JOURNAL Cloning of three CD20 homolog from human, putative calcium channels
AUTHORS Unpublished
TITLE 3 (bases 1 to 691)
REFERENCE Ishibashi, K.
JOURNAL Direct Submission
AUTHORS Submitted (20-APR-1998) Kenichi Ishibashi, Tokyo Medical and Dental
TITLE University, 2nd Internal Medicine; Yushima 1-5-45, Bunkyo, Tokyo
JOURNAL 113-8519, Japan (E-mail: kishibashi.med2med.tmd.ac.jp,
Tel:81-3-5803-5223, Fax:81-3-5803-0132)
FEATURES
source
location/Qualifiers
1..691
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="testis"
52..654
/gene="MS4A5"
52..654
/gene="MS4A5"
52..654
/codon_start=1
/product="MS4A5"
/protein_id="BAB18739.1"
/db_xref="GI:11559214"
/translaction="MDSSTASHPVFLVPPPEITASEVSESTELSATTFSTOSPLQKLF
RKMILGIIQILFGIMFSPGVIFLTLKPYPRPFILSGPFGMSVLPINSAGFL
IAVRKTEETLIIISRIIMNLSAIGALIGIILTFGLDNIYICGYSHNSQCKAATV
VLFGLITLMTFSIETLFIPLPSIILCHSEDDCEQCC"
BASE COUNT 195 a 142 c 117 g 237 t
ORIGIN
Query Match 77.2%; Score 347.4; DB 9; Length 691;
Best Local Similarity 97.0%; Pred. No. 2.5e-67;
Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 atgagattcaagcagcagacagtcgcgtgttctcgtgtaatttcctccagaaatcactgct 60
Db 52 ATGGATTCACGACGACGACAGTCGCGGTTCGTGGTATTTCCGCCGAATTCACACTGCT 111
QY 61 tcagatatagatccacagaaacttcacgcacagaccttttaactcaaaagcccttgcaa 120
Db 112 TCAGATATAGATCCACAGAACTTTTCAGCAGACCTTTTCAACTCAAGCCCTTGCAA 171

```

```

QY 121 aaattatgctagaaaataatccttagggactaccagatccctggttggaattatg 180
Db 172 AATATTTTTCGTGAGAAAATGAAAATCTTAGGACATATCCAGATCTGTTTGGAAATTAATG 231
QY 181 acccttccttgaggatlaacctccttcaccttgctlaaaccatalccaaagttccc 240
Db 232 ACCCTTCTTTTGAGATATATCTCCCTTTTCACCTTGTTAAACCATATCCAGGTTTCCC 291
QY 241 ttataatctcttcagagatataccatctcgggctcgtgtttgttaattaatctgagcc 300
Db 292 TTTATATTTCTTTCAGAGATATTCATCTGGGCTCTGTTTGTCATTAATCTGGAGCC 351
QY 301 ttcccaatgcagtgaaagaaacacagaaactctcgggaatttggttcacatgag 360
Db 352 TTCCTAATTTGACGTAGAAAAGAAACACAGAACTCTGATATATTTGAGCCGAATTAATG 411
QY 361 acttc 365
Db 412 AATTT 416

RESULT 5
LOCUS AF237907 694 bp mRNA PRI 17-APR-2001
DEFINITION Homo sapiens MS4A5 protein mRNA, complete cds.
ACCESSION AF237907
VERSION AF237907.1 GI:13649400
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 694)
AUTHORS Liang, Y. and Tedder, T.F.
TITLE Identification of a cd20-, fcepsilonibeta-, and htm4-related gene
JOURNAL family: sixteen new ms4a family members expressed in human and
AUTHORS mouse
JOURNAL Genomics 72 (2), 119-127 (2001)
MEDLINE 21295030
REFERENCE 2 (bases 1 to 694)
AUTHORS Liang, Y. and Tedder, T.F.
JOURNAL Direct Submission
AUTHORS Submitted (24-FEB-2000) Immunology, Duke Medical Center, Research
JOURNAL Dr., Durham, NC 27710, USA
FEATURES
source
location/Qualifiers
1..694
/organism="Homo sapiens"
/db_xref="taxon:9606"
55..657
/codon_start=1
/product="MS4A5 protein"
/protein_id="AAK37416.1"
/db_xref="GI:13649401"
/translaction="MDSSTASHPVFLVPPPEITASEVSESTELSATTFSTOSPLQKLF
RKMILGIIQILFGIMFSPGVIFLTLKPYPRPFILSGPFGMSVLPINSAGFL
IAVRKTEETLIIISRIIMNLSAIGALIGIILTFGLDNIYICGYSHNSQCKAATV
VLFGLITLMTFSIETLFIPLPSIILCHSEDDCEQCC"
BASE COUNT 198 a 143 c 116 g 237 t
ORIGIN
Query Match 77.2%; Score 347.4; DB 9; Length 694;
Best Local Similarity 97.0%; Pred. No. 2.5e-67;
Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 atgagattcaagcagcagacagtcgcgtgttctcgtgtaatttcctccagaaatcactgct 60
Db 55 ATGGATTCACGACGACGACAGTCGCGGTTCGTGGTATTTCCGCCGAATTCACACTGCT 114
QY 61 tcagatatagatccacagaaacttcacgcacagaccttttaactcaaaagcccttgcaa 120
Db 115 TCAGATATAGATCCACAGAACTTTTCAGCAGACCTTTTCAACTCAAGCCCTTGCAA 174

```

QY	121	aaattttctgcataaataaagaacatttaggacatccagaatcctgtttggaattatg	180
Db	175	AAATTATTTGGTGAAGAAAANGAATAACTTAGAGCACTATCCAGATCCTGTTTGGAAATTATG	234
QY	181	acctttctttggaggtatctctcctttcaacctgtttaaaccaatccaaggtttccc	240
Db	235	ACCTTTCTTTTGGAGTATTCCTTCCTTTTCACCTTTGTATAAACCATATCCAAAGTTTCCC	294
QY	241	tttatattctttcaagatatcatctctgggggtctcgttttftcattatcttggaagcc	300
Db	295	TTTTATATTTCTTTTCAGAGATATCCATTCTG6GGCTGTGTTTGTCTATTAATCTTGAGAGCC	354
QY	301	ttcctaattgcagtgagaaagaaacacagaagaactctgggaatttgatttcatgatg	360
Db	355	TTTCCTAATTTGCAGTGAAGAAAAGAAAACACAGAAACTCTGATTAATTATGACCGCAATATG	414
QY	361	accttt 365	
Db	415	AAATTT 419	
RESULT	6		
AF321127		697 bp	PRI 07-FEB-2001
LOCUS	AF321127	697 bp	mRNA
DEFINITION	Homo sapiens testis-expressed transmembrane-4 protein (TETM4)		mRNA.
ACCESSION	AF321127		complete cds.
VERSION	AF321127.1	GI:12698681	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 697)		
TITLE	Hulett,M.D., Paglier,E., Hornby,J.R., Hogarth,P.M., Eyre,H.J., Baker,E., Crawford,J., Sutherland,G.R., Ohms,S.J. and Parish,C.R.		
JOURNAL	Isolation, tissue distribution, and chromosomal localization of a novel testis-specific human four-transmembrane gene related to CD20		
MEDLINE	and Fcgspl10RI-Beta		
REFERENCE	Biochem. Biophys. Res. Commun. 280 (1), 374-379 (2001)		
AUTHORS	2 (bases 1 to 697)		
TITLE	Hulett,M.D.		
JOURNAL	Direct Submission		
MEDLINE	Submitted (14-NOV-2000) Division of Immunology and Cell Biology,		
REFERENCE	John Curtin School of Medical Research, Mills Road, Canberra, ACT		
AUTHORS	2601, Australia		
TURES	Location/Qualifiers		
SOURCE	1..697		
gene	/organism="Homo sapiens"		
CDs	/db_xref="taxon:9606"		
	/chromosome="11"		
	/map="11q12"		
	1..697		
	/gene="TETM4"		
	/gene="TETM4"		
	/protein_id="AAK01641.1"		
	/product="testis-expressed transmembrane-4 protein"		
	/codon_start=1		
	/protein_id="AAK01641.1"		
	/db_xref="GI:12698682"		
	/translation="MDSSTAHSPVFLVPEPEITASEYSETELSTTFESTOSPLOKLFAR		
	KKMLCTGICILGIMTFSEGVIFLPTLLKPYPEPIPLFGSGYPMGSVYFINSQATL		
	IAYRKTKTEILISRLIMFSLNMFSLAGATAGIILITLFGILLDONTYICGISHONSQCKAYT		
	VLLGILITLMTPTIILFTLSIPSLTSGHSEDCDEQCC"		
BASE COUNT	199 a 142 c 119 g 237 t		
ORIGIN			

QY	1	atgattcaagacacgcagatccgggtgttcctggtattctctccagaataactgct	60
Db	57	ATGATTCGAAGACGCCACACACTCCGGTGTTCTTGTATTTCTCCAGAAATCAGTCT	116
QY	61	tcagaatatgagtcacagaacttcagccagacctttcaactcaagcccttgcaa	120
Db	117	TCGAAATATGAGTCCACAGAACTTTCAGCCACGACCTTTCAACTCAAAAGCCCTTGCAA	176
QY	121	aaatatttgctagaaaatgaaaacttttagggacatccagatcctglttggaaattatg	180
Db	177	AAATATTTGCTAGAAAATGAAAATCTTAGGGACATTCAGATCCGTGTTGGAAATTAATG	236
QY	181	acctttcctttggagtattctctcttttcaacttggttttaaaccatatccagaagttccc	240
Db	237	ACCTTTCTTTTGGATTTATCTTCCTTTACACTTTGTATAAACATATCCAAAGTTTCCC	296
QY	241	tttataattctcttcagagataccatctctggygcctcgtttgtgttcataatctggagcc	300
Db	297	TTTATATTTCTTTCAGGAATATCCATCTCTGGGGCTCTGTTTGTTCATTTATCTGGAGCC	356
QY	301	ttccctaattgagtgtaaaaagaaaaccacagaagaacctctgggaattttgattacattgatg	360
Db	357	TTTCCTTATTTGCAGTGAAAAGAAAACCAAGAAAATCTGTATATTTAGGCCGAATATATG	416
QY	361	acctt 365	
Db	417	AATTT 421	
RESULT	7		
LOCUS	AX179785	417 bp	DNA
DEFINITION	Sequence 5 from Patent WO0146417.		PAT
ACCESSION	AX179785		06-AUG-2001
VERSION	AX179785.1	GI:15132146	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 417)		
JOURNAL	Walke,D.W. and Turner,C.A.		
FEATURES	Human membrane proteins and polynucleotides encoding the same		
Source	having homology to cd20 proteins and ige receptors		
	Patent: WO 0146417-A 5 28-JUN-2001;		
	Lexicon Genetics Incorporated (US)		
	Location/Qualifiers		
	1..417		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	115 a	91 c	69 g
ORIGIN			142 t
Query Match	75.7%;	Score 340.8;	DB 6; Length 417;
Best Local Similarity	94.1%;	Pred. No. 7.2e-66;	
Matches 354;	Conservative 0;	Mismatches 22;	Indels 0; Gaps 0;

Db 181 ACCTTTCTTTGGAGTATCTCTCTTTACCTGTGTAACCATATATCCAGTTCC 240
Qy 241 titatatcttcttcaagatcatctcgtggcgtctgtttcttcaattatctgaacc 300
Db 241 TTTATATTTCTTCAGAGATATCCATTCTGGGCTGTGTTGTCATTAATCTTGAGCC 300
Qy 301 ttccaatgcagtgaaagaaacacagaacactcggaaatttgattacatgacg 360
Db 301 TTTCCTAATTCGACGTGAAGAAAGAAACACAGAAACTGTGATCAAACTACATTGTGTT 360
Qy 361 acttcacgattatg 376
Db 361 ATTCTCACCATAATAG 376

RESULT 8
AC027787 138097 bp DNA HTG 18-AUG-2000
LOCUS Homo sapiens chromosome 15 clone RP11-196E16 map 15, LOW-PASS
DEFINITION
SEQUENCE SAMPLING.
AC027787
AC027787.2 GI:9845160
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 138097)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
JOURNAL Homo sapiens chromosome 15, clone RP11-196E16
REFERENCE 2 (bases 1 to 138097)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beka, F.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J. S.,
Dodge, S., Domingo, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Glade, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, C., Hagos, B., Heaford, A., Horton, L.,
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,
Meldrum, J., Menus, L., Mihova, T., Miranda, C., Mlenka, Y., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT
Submitted (01-Apr-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 18, 2000 this sequence version replaced gi:7382629.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L9008
Center clone name: 196_E_16

* NOTE: This record contains 163 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely

* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1 663: contig of 663 bp in length
664 763: gap of 100 bp
764 1464: contig of 701 bp in length
1465 1564: gap of 100 bp
1565 2268: contig of 704 bp in length
2269 2368: gap of 100 bp
2369 3042: contig of 674 bp in length
3043 3142: gap of 100 bp
3143 3620: contig of 678 bp in length
3621 3920: gap of 100 bp
3921 4626: contig of 706 bp in length
4627 4726: gap of 100 bp
4727 5423: contig of 697 bp in length
5424 5523: gap of 100 bp
5524 6234: contig of 711 bp in length
6235 6334: gap of 100 bp
6335 7040: contig of 706 bp in length
7041 7140: gap of 100 bp
7141 7637: contig of 697 bp in length
7638 7937: gap of 100 bp
7938 8632: contig of 695 bp in length
8633 8732: gap of 100 bp
8733 9430: contig of 698 bp in length
9431 9530: gap of 100 bp
9531 10214: contig of 684 bp in length
10215 10314: gap of 100 bp
10315 10980: contig of 666 bp in length
10981 11080: gap of 100 bp
11081 11783: contig of 703 bp in length
11784 11883: gap of 100 bp
11884 12582: contig of 699 bp in length
12583 12682: gap of 100 bp
12683 13389: contig of 707 bp in length
13390 13489: gap of 100 bp
13490 14186: contig of 697 bp in length
14187 14286: gap of 100 bp
14287 14996: contig of 710 bp in length
14997 15096: gap of 100 bp
15097 15796: contig of 700 bp in length
15797 15896: gap of 100 bp
15897 16583: contig of 687 bp in length
16584 16683: gap of 100 bp
16684 17376: contig of 693 bp in length
17377 17476: gap of 100 bp
17477 18178: contig of 702 bp in length
18179 18278: gap of 100 bp
18279 18962: contig of 684 bp in length
18963 19062: gap of 100 bp
19063 19673: contig of 611 bp in length
19674 19773: gap of 100 bp
19774 20436: contig of 663 bp in length
20437 20536: gap of 100 bp
20537 21233: contig of 697 bp in length
21234 21333: gap of 100 bp
21334 22042: contig of 709 bp in length
22043 22142: gap of 100 bp
22143 22808: contig of 666 bp in length
22809 22908: gap of 100 bp
22909 23595: contig of 687 bp in length
23596 23695: gap of 100 bp
23696 24391: contig of 696 bp in length
24392 24491: gap of 100 bp
24492 25190: contig of 699 bp in length
25191 25290: gap of 100 bp
25291 25981: contig of 691 bp in length
25982 26081: gap of 100 bp

```

* 26082 26765: contig of 684 bp in length
* 26766 26865: gap of 100 bp
* 26866 27563: contig of 698 bp in length
* 27564 27663: gap of 100 bp
* 27664 28377: contig of 714 bp in length
* 28378 28477: gap of 100 bp
* 28478 29184: contig of 707 bp in length
* 29185 29284: gap of 100 bp
* 29285 29983: contig of 699 bp in length
* 29984 30083: gap of 100 bp
* 30084 30775: contig of 692 bp in length
* 30776 30875: gap of 100 bp
* 30876 31565: contig of 680 bp in length
* 31566 31663: gap of 100 bp
* 31666 32360: contig of 695 bp in length
* 32361 32460: gap of 100 bp
* 32461 33157: contig of 697 bp in length
* 33158 33257: gap of 100 bp
* 33258 33951: contig of 694 bp in length
* 33952 34051: gap of 100 bp
* 34052 34735: contig of 684 bp in length
* 34736 34835: gap of 100 bp
* 34836 35483: contig of 648 bp in length
* 35484 35583: gap of 100 bp
* 35584 36308: contig of 725 bp in length
* 36309 36408: gap of 100 bp
* 36409 37108: contig of 700 bp in length
* 37109 37208: gap of 100 bp
* 37209 37892: contig of 684 bp in length
* 37893 37992: gap of 100 bp
* 37993 38672: contig of 680 bp in length
* 38673 38772: gap of 100 bp
* 38773 39473: contig of 703 bp in length
* 39476 39575: gap of 100 bp
* 39576 40290: contig of 715 bp in length
* 40291 40390: gap of 100 bp
* 40391 41087: contig of 697 bp in length
* 41088 41187: gap of 100 bp
* 41188 41891: contig of 704 bp in length
* 41892 41991: gap of 100 bp
* 41992 42679: contig of 688 bp in length
* 42680 42779: gap of 100 bp
* 42780 43500: contig of 721 bp in length
* 43501 43600: gap of 100 bp
* 43601 44295: contig of 695 bp in length
* 44296 44395: gap of 100 bp
* 44396 45088: contig of 693 bp in length
* 45089 45188: gap of 100 bp
* 45189 45880: contig of 692 bp in length
* 45881 45980: gap of 100 bp
* 45981 46683: contig of 703 bp in length
* 46684 46783: gap of 100 bp
* 46784 47481: contig of 698 bp in length
* 47482 47581: gap of 100 bp
* 47582 48290: contig of 709 bp in length
* 48291 48390: gap of 100 bp
* 48391 49077: contig of 687 bp in length
* 49078 49177: gap of 100 bp
* 49178 49879: contig of 702 bp in length
* 49880 49979: gap of 100 bp
* 49980 50684: contig of 705 bp in length
* 50685 50784: gap of 100 bp
* 50785 51507: contig of 723 bp in length
* 51508 51607: gap of 100 bp
* 51608 52320: contig of 713 bp in length
* 52321 52420: gap of 100 bp
* 52421 53133: contig of 713 bp in length
* 53134 53233: gap of 100 bp
* 53234 53933: contig of 700 bp in length
* 53934 54033: gap of 100 bp
* 54034 54703: contig of 670 bp in length

```

Query Match

34.1%: Score 153.4; DB 2; Length 138097;

```

Best Local Similarity 96.3%; Pred. No. 2.8e-24;
Matches 157; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 atggattcaagcaccgacgagtcggtgtgttcggtatcccccgaactactgct 60
|||||
DB 109596 ATGGATTCAAGCACCACGACAGTCGGGTGTGTGTTTCGTCACGAATCACTCT 109655

QY 61 tcagatatgacgcacgaacttcacgcacgaccttcacccaagaacccctgcga 120
|||||
DB 109656 TCAGATATGATGATCCACGAACTTTCAGCGACGACTTTTCACATCAAAAGCCCTTGC 109715

QY 121 aaattatttcgtagaaaaatgaatcttagggactatccaga 163
|||||
DB 109716 AAATTATTTCGTAGAAAAATGAATCTTAGGGCTTAGTAAGA 109758

RESULT 9
LOCUS AC027787/c 138097 bp DNA HTG 18-AUG-2000
DEFINITION Homo sapiens chromosome 15 clone RP11-196E16 map 15, LOW-PASS
SEQUENCE SAMPLING.
ACCESSION AC027787
VERSION AC027787.2 GI:9845160
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 138097)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 15, clone RP11-196E16
Unpublished
2 (bases 1 to 138097)
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bada, F.,
Boguslavsky, L., Bouhassira, B., Brown, A., Burnett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collins, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J. S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardina, S., Ginde, S., Goette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Lacombe, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Lien, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPeeters, R.,
Meldrum, J., Menus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neil, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rotman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessier, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 18, 2000 this sequence version replaced gi:7382629.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: U9008
Center clone name: 196_E_16
-----
* NOTE: This record contains 163 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads

```


* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
* 663: contig of 663 bp in length
* 664 763: gap of 100 bp
* 764 1464: contig of 701 bp in length
* 1465 1564: gap of 100 bp
* 1565 2268: contig of 704 bp in length
* 2269 2368: gap of 100 bp
* 2369 3042: contig of 674 bp in length
* 3043 3142: gap of 100 bp
* 3143 3820: contig of 678 bp in length
* 3821 3920: gap of 100 bp
* 3921 4626: contig of 706 bp in length
* 4627 4726: gap of 100 bp
* 4727 5423: contig of 697 bp in length
* 5424 5523: gap of 100 bp
* 5524 6234: contig of 711 bp in length
* 6235 6334: gap of 100 bp
* 6335 7040: contig of 706 bp in length
* 7041 7140: gap of 100 bp
* 7141 7837: contig of 697 bp in length
* 7838 7937: gap of 100 bp
* 7938 8632: contig of 695 bp in length
* 8633 8732: gap of 100 bp
* 8733 9430: contig of 698 bp in length
* 9431 9530: gap of 100 bp
* 9531 10214: contig of 684 bp in length
* 10215 10314: gap of 100 bp
* 10315 10980: contig of 666 bp in length
* 10981 11080: gap of 100 bp
* 11081 11783: contig of 703 bp in length
* 11784 11883: gap of 100 bp
* 11884 12582: contig of 699 bp in length
* 12583 12682: gap of 100 bp
* 12683 13389: contig of 707 bp in length
* 13390 13489: gap of 100 bp
* 13490 14186: contig of 697 bp in length
* 14187 14286: gap of 100 bp
* 14287 14996: contig of 710 bp in length
* 14997 15096: gap of 100 bp
* 15097 15796: contig of 700 bp in length
* 15797 15896: gap of 100 bp
* 15897 16583: contig of 687 bp in length
* 16584 16683: gap of 100 bp
* 16684 17376: contig of 693 bp in length
* 17377 17476: gap of 100 bp
* 17477 18178: contig of 702 bp in length
* 18179 18278: gap of 100 bp
* 18279 18962: contig of 684 bp in length
* 18963 19062: gap of 100 bp
* 19063 19673: contig of 611 bp in length
* 19674 19773: gap of 100 bp
* 19774 20436: contig of 663 bp in length
* 20437 20536: gap of 100 bp
* 20537 21233: contig of 697 bp in length
* 21234 21333: gap of 100 bp
* 21334 22042: contig of 709 bp in length
* 22043 22142: gap of 100 bp
* 22143 22808: contig of 666 bp in length
* 22809 22908: gap of 100 bp
* 22909 23595: contig of 687 bp in length
* 23596 23695: gap of 100 bp
* 23696 24391: contig of 696 bp in length
* 24392 24491: gap of 100 bp
* 24492 25190: contig of 699 bp in length
* 25191 25290: gap of 100 bp
* 25291 25981: contig of 691 bp in length

25982 26081: gap of 100 bp
* 26082 26765: contig of 684 bp in length
* 26766 26865: gap of 100 bp
* 26866 27563: contig of 698 bp in length
* 27564 27663: gap of 100 bp
* 27664 28377: contig of 714 bp in length
* 28378 28477: gap of 100 bp
* 28478 29184: contig of 707 bp in length
* 29185 29284: gap of 100 bp
* 29285 29983: contig of 699 bp in length
* 29984 30083: gap of 100 bp
* 30084 30775: contig of 692 bp in length
* 30776 30875: gap of 100 bp
* 30876 31565: contig of 690 bp in length
* 31566 31665: gap of 100 bp
* 31666 32360: contig of 695 bp in length
* 32361 32460: gap of 100 bp
* 32461 33157: contig of 697 bp in length
* 33158 33257: gap of 100 bp
* 33258 33951: contig of 694 bp in length
* 33952 34051: gap of 100 bp
* 34052 34735: contig of 684 bp in length
* 34736 34835: gap of 100 bp
* 34836 35483: contig of 648 bp in length
* 35484 35583: gap of 100 bp
* 35584 36308: contig of 725 bp in length
* 36309 36408: gap of 100 bp
* 36409 37108: contig of 700 bp in length
* 37109 37208: gap of 100 bp
* 37209 37892: contig of 684 bp in length
* 37893 37992: gap of 100 bp
* 37993 38672: contig of 680 bp in length
* 38673 38772: gap of 100 bp
* 38773 39473: contig of 703 bp in length
* 39476 39575: gap of 100 bp
* 39576 40290: contig of 715 bp in length
* 40291 40390: gap of 100 bp
* 40391 41087: contig of 697 bp in length
* 41088 41187: gap of 100 bp
* 41188 41891: contig of 704 bp in length
* 41892 41991: gap of 100 bp
* 41992 42679: contig of 688 bp in length
* 42680 42779: gap of 100 bp
* 42780 43500: contig of 721 bp in length
* 43501 43600: gap of 100 bp
* 43601 44295: contig of 695 bp in length
* 44296 44395: gap of 100 bp
* 44396 45088: contig of 693 bp in length
* 45089 45188: gap of 100 bp
* 45189 45880: contig of 692 bp in length
* 45881 45980: gap of 100 bp
* 45981 46683: contig of 703 bp in length
* 46684 46783: gap of 100 bp
* 46784 47481: contig of 698 bp in length
* 47482 47581: gap of 100 bp
* 47582 48290: contig of 709 bp in length
* 48291 48390: gap of 100 bp
* 48391 49077: contig of 687 bp in length
* 49078 49177: gap of 100 bp
* 49178 49879: contig of 702 bp in length
* 49880 49979: gap of 100 bp
* 49980 50684: contig of 705 bp in length
* 50685 50784: gap of 100 bp
* 50785 51507: contig of 723 bp in length
* 51508 51607: gap of 100 bp
* 51608 52320: contig of 713 bp in length
* 52321 52420: gap of 100 bp
* 52421 53133: contig of 713 bp in length
* 53134 53233: gap of 100 bp
* 53234 53933: contig of 700 bp in length
* 53934 54033: gap of 100 bp
* 54034 54703: contig of 670 bp in length

AUTHORS
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

TITLE
Homo sapiens 161,039 genomic DNA of 11g

JOURNAL
Published Only in Database (2001) In press

REFERENCE
2 (bases 1 to 161039)

AUTHORS
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

TITLE
Direct Submission

JOURNAL
Submitted (26-JAN-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suhiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.go.jp, URL: http://hnp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)

COMMENT
Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hnp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
Project Information
Center project name: Humdraft1
Center clone name: Rp11-71066
Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 157456 bases at least Q40
Consensus quality: 156960 bases at least Q30
Consensus quality: 159457 bases at least Q20
Insert size: 158639; sum-of-contigs
Quality coverage: 8.61x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
15 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.

1 42322 contig of 42322 bp in length
42423 63418 contig of 20996 bp in length
63519 77969 contig of 14451 bp in length
78070 93182 contig of 15113 bp in length
93283 104093 contig of 10811 bp in length
104194 115702 contig of 11509 bp in length
115803 125183 contig of 9381 bp in length
125284 132558 contig of 7275 bp in length
132659 137219 contig of 4561 bp in length
137320 143276 contig of 5957 bp in length
143377 148922 contig of 5546 bp in length
149023 154159 contig of 5137 bp in length
154260 157686 contig of 3427 bp in length
157787 159234 contig of 1448 bp in length
159335 161039 contig of 1705 bp in length.

NOTE: This is a 'working draft' sequence. It currently
consists of 15 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 42322: contig of 42322 bp in length
* 42323 42422: gap of 100 bp
* 42423 63418: contig of 20996 bp in length
* 63419 63518: gap of 100 bp
* 63519 77969: contig of 14451 bp in length
* 77970 78069: gap of 100 bp
* 78070 93182: contig of 15113 bp in length
* 93183 93282: gap of 100 bp
* 93283 104093: contig of 10811 bp in length
* 104094 104193: gap of 100 bp
* 104194 115702: contig of 11509 bp in length
* 115703 115802: gap of 100 bp

115803 125183: contig of 9381 bp in length
* 125184 125283: gap of 100 bp
* 125284 132558: contig of 7275 bp in length
* 132559 132658: gap of 100 bp
* 132659 137219: contig of 4561 bp in length
* 137220 137319: gap of 100 bp
* 137320 143276: contig of 5957 bp in length
* 143277 143376: gap of 100 bp
* 143377 148922: contig of 5546 bp in length
* 148923 149022: gap of 100 bp
* 149023 154159: contig of 5137 bp in length
* 154160 154259: gap of 100 bp
* 154260 157686: contig of 3427 bp in length
* 157687 157786: gap of 100 bp
* 157787 159234: contig of 1448 bp in length
* 159235 159334: gap of 100 bp
* 159335 161039: contig of 1705 bp in length.

FEATURES
source
1. 161039
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11g"
/clone="Rp11-71066"
1. 42322
/note="assembly_fragment"
42423..63418
/note="assembly_fragment"
63519..77969
/note="assembly_fragment"
78070..93182
/note="assembly_fragment"
93283..104093
/note="assembly_fragment"
104194..115702
/note="assembly_fragment"
115803..125183
/note="assembly_fragment"
125284..132558
/note="assembly_fragment"
132659..137219
/note="assembly_fragment"
137320..143276
/note="assembly_fragment"
143377..148922
/note="assembly_fragment clone_end:T7 vector_side:left"
149023..154159
/note="assembly_fragment"
154260..157686
/note="assembly_fragment"
157787..159234
/note="assembly_fragment"
159335..161039
/note="assembly_fragment"

BASE COUNT 49155 a 31166 c 31523 g 47795 t 1400 others
ORIGIN

Query Match 34.1%; Score 153.4; DB 2; Length 161039;
Best Local Similarity 96.3%; Pred. No. 2.8e-24;
Matches 157; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 atgattcaagcaccgcacacagtcgcgtgtcttcgatttcctccagaatcactgct 60
|||||
DB 149618 ATGGATTCAGACACCGCACAGTCGCGGTTCGTRATTTCTCCAGAAATCACTGCT 149559
|||||

OY 61 tcagaatatgagtcacagaaacttcagccacagcactttcaactcaagaagcccttgcaa 120
|||||
DB 149558 TCAGAAATATGATGATCCACAGAACTTCACGACGACCTTTCAACTCAAGCCCTTGCAA 149499
|||||

OY 121 aatatattgctcagaanaaatgaaatcttaggaactatccaga 163
|||||
DB 149498 AATATATTGCTGTAAGAAATGAATAATCTTAGCGGTACTAAGA 149456
|||||

RESULT 12
AP001034/c
LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP11-729B4 map 11q12, WORKING
DRAFT SEQUENCE, 19 unordered pieces.
ACCESSION AP001034 GI:11176992
VERSION AP001034.4
KEYWORDS HTG: HTGS-PHASE1; HTGS-DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-729B4.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 166804)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 166,804 genomic DNA of 11q12
Published Only in Database (2000) In press
2 (bases 1 to 166804)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (06-JAN-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Nov 14, 2000 this sequence version replaced gi:8117704.
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: Humrap11
Center clone name: RP11-729B4
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 162254 bases at least Q40
Consensus quality: 163858 bases at least Q30
Consensus quality: 164544 bases at least Q20
Insert size: 165004; sum-of-contigs
Quality coverage: 9.28x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
19 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 22398 contig of 22398 bp in length
22499 45238 contig of 22740 bp in length
45339 45238 contig of 18940 bp in length
64379 77579 contig of 13201 bp in length
77680 93718 contig of 16039 bp in length
93819 108057 contig of 14239 bp in length
108158 117107 contig of 8849 bp in length
117107 126302 contig of 9196 bp in length
126403 133479 contig of 7077 bp in length
13380 140425 contig of 6846 bp in length
140526 144653 contig of 4128 bp in length
144754 149455 contig of 4601 bp in length
149455 154522 contig of 5068 bp in length
154522 155919 contig of 1297 bp in length
155919 156019 gap of 100 bp
156019 159504 contig of 3485 bp in length
159504 159604 gap of 100 bp
159604 159679 contig of 75 bp in length
159679 162418 contig of 100 bp
162418 162518 gap of 2639 bp in length
162518 165383 contig of 2865 bp in length
165383 165483 gap of 100 bp
165483 166804 contig of 1321 bp in length.

NOTE: This is a 'working draft' sequence. It currently
consists of 19 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 22398: contig of 22398 bp in length
22399 22498: gap of 100 bp
22499 45238: contig of 22740 bp in length
45239 45338: gap of 100 bp
45339 64278: contig of 18940 bp in length
64279 64378: gap of 100 bp
64379 77579: contig of 13201 bp in length
77580 77679: gap of 100 bp
77680 93718: contig of 16039 bp in length
93719 93818: gap of 100 bp
93819 108057: contig of 14239 bp in length
108058 108157: gap of 100 bp
108158 117006: contig of 8849 bp in length
117007 117106: gap of 100 bp
117107 126302: contig of 9196 bp in length
126303 126402: gap of 100 bp
126403 133479: contig of 7077 bp in length
133480 133579: gap of 100 bp
133580 140425: contig of 6846 bp in length
140426 140525: gap of 100 bp
140526 144653: contig of 4128 bp in length
144654 144753: gap of 100 bp
144754 149354: contig of 4601 bp in length
149355 149454: gap of 100 bp
149455 154522: contig of 5068 bp in length
154523 154622: gap of 100 bp
154623 155919: contig of 1297 bp in length
155920 156019: gap of 100 bp
156020 159504: contig of 3485 bp in length
159505 159604: gap of 100 bp
159605 159679: contig of 75 bp in length
159680 159779: gap of 100 bp
159780 162418: contig of 2639 bp in length
162419 162518: gap of 100 bp
162519 165383: contig of 2865 bp in length
165384 165483: gap of 100 bp
165484 166804: contig of 1321 bp in length.
Location/Qualifiers
1..166804
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q12"
/clone="RP11-729B4"
1..22398
/note="assembly_fragment"
22499..45238
/note="assembly_fragment"
45339..64278
/note="assembly_fragment"
64379..77579
/note="assembly_fragment"
77680..93718
/note="assembly_fragment"
93819..108057
/note="assembly_fragment"
108158..117006
/note="assembly_fragment"
117107..126302
/note="assembly_fragment"
126403..133479
/note="assembly_fragment"
133580..140425
/note="assembly_fragment"
140526..144653

```

misc_feature /note="assembly_fragment"
144754 .149354
misc_feature /note="assembly_fragment"
149455 .154322
misc_feature /note="assembly_fragment"
154623 .155919
misc_feature /note="assembly_fragment clone_end:SP6 vector_side:right"
156020 .159504
misc_feature /note="assembly_fragment"
159605 .159679
misc_feature /note="assembly_fragment clone_end:T7 vector_side:left"
159780 .162418
misc_feature /note="assembly_fragment"
162519 .165383
misc_feature /note="assembly_fragment"
165484 .166804
misc_feature /note="assembly_fragment"
BASE COUNT 50706 a 31523 c 32296 g 50478 t 1801 others
ORIGIN

```

```

Query Match 34.1%; Score 153.4; DB 2; Length 166804;
Best Local Similarity 96.3%; Pred. No. 2.8e-24;
Matches 157; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

OY 1 atggattcaagcagcagcagcagtcggtgttcttgattccctccagaatacactgct 60
DB 11755 ATGGATTCAAGCAGCAGCAGCAGTCGGTGTTCGTTCGTTCTTCCTCCAGAAATCACTGCT 11696
OY 61 taagaataagatccacagaactttcagccagcagcctttcaactcaagcccttgcaa 120
DB 11695 TCAGAAATAGATGCCACAGAACTTTCAGCCAGCAGCCTTTCAACTCAAGCCCTTGCAA 11636
OY 121 aaattattgctagaaaaatgaataatcttagggactatccaga 163
DB 11635 AAATTATTGCTAGAAAAATGAATAATCTTAGGGGTAAGTACA 11593

```

```

RESULT 13
AC090401/C 166820 bp DNA HTG 21-JUL-2001
LOCUS Homo sapiens chromosome 11 clone RP11-729B4 map 11, WORKING DRAFT
DEFINITION
SEQUENCE, 2 ordered pieces.
AC090401
AC090401.2 GI:14971278
VERSION
KEYWORDS
SOURCE
HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
AUTHORS
1 (bases 1 to 166820)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 11, clone RP11-729B4
Unpublished
2 (bases 1 to 166820)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barra, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A.,
Camarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
Gardina, S., Ginde, S., Goyette, M., Graham, L., Grande-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T.,
Lehoczky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
McPheeters, R., Meldrum, J., Menus, L., Mihova, T., Mienga, V.,
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,
O'Connor, F., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pletre, N., Pollara, V., Raymond, C., Retta, R.,
Rienack, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,
Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P.,
Soune, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,

```

TITLE
JOURNAL
COMMENT

Travers, M., Travis, N., Trigilio, J., Vassiliou, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 20, 2001 this sequence version replaced gi:12958037.
ALL repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RW/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L12709
Center clone name: 729_B_4

----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 166378 bases at least Q40
Consensus quality: 166639 bases at least Q30
Consensus quality: 166692 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 166720; sum-of-contigs
Quality coverage: 12.2 in Q20 bases; agarose-fp
Quality coverage: 12.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 1096 1095: contig of 1095 bp in length
* 1096 1195: gap of 100 bp
* 1196 166820: contig of 165625 bp in length.
Location/Qualifiers
1. 166820
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-729B4"
/clone_1lb="RP11-729B4 Human Male BAC"
1. 1095
/note="assembly_fragment"
1196. 166820
/note="assembly_fragment"
clone_end:T7
vector_side:right"

```

BASE COUNT 46436 a 31826 c 32895 g 55563 t 100 others
ORIGIN

```

```

Query Match 34.1%; Score 153.4; DB 2; Length 166804;
Best Local Similarity 96.3%; Pred. No. 2.8e-24;
Matches 157; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

OY 1 atggattcaagcagcagcagcagtcggtgttcttgattccctccagaatacactgct 60
DB 144100 ATGGATTCAAGCAGCAGCAGCAGTCGGTGTTCGTTCGTTCTTCCTCCAGAAATCACTGCT 144041
OY 61 taagaataagatccacagaactttcagccagcagcctttcaactcaagcccttgcaa 120
DB 144040 TCAGAAATAGATGCCACAGAACTTTCAGCCAGCAGCCTTTCAACTCAAGCCCTTGCAA 143981
OY 121 aaattattgctagaaaaatgaataatcttagggactatccaga 163

```

DB 143980 AATATTGCTAGAAAATGAAATCTTAGGGGTAAGTAA 143938

RESULT 14

AP003127

LOCUS

DEFINITION Homo sapiens chromosome 11 clone RP11-71066 map 11q, WORKING DRAFT

SEQUENCE, 15 unordered pieces.

AP003127

AP003127.1 GI:12597183

HTG: HTGS.PHASE1.DRAFT.

SOURCE Homo sapiens DNA, clone:RP11-71066.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 161039)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

2 (bases 1 to 161039)

Published Only in Database (2001) In press

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Submitted (26-JAN-2001) Masahira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,

Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT

Center: RIKEN Genomic Sciences Center(GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@gsc.riken.go.jp

Project Information

Center project name: Hummap11

Center clone name: RP11-71066

Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 157456 bases at least Q40

Consensus quality: 158960 bases at least Q30

Insert size: 159639; sum-of-contigs

Quality coverage: 8.61x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 42322 contig of 42322 bp in length
42423 63418 contig of 20996 bp in length
63519 77969 contig of 14451 bp in length
78070 93182 contig of 15113 bp in length
93283 104093 contig of 10811 bp in length
104194 115702 contig of 11509 bp in length
115803 125183 contig of 9381 bp in length
125284 132558 contig of 7275 bp in length
132659 137219 contig of 4561 bp in length
137320 143276 contig of 5957 bp in length
143377 148922 contig of 5546 bp in length
149023 154159 contig of 5137 bp in length
154260 157686 contig of 3427 bp in length
157787 159234 contig of 1448 bp in length
159335 161039 contig of 1705 bp in length.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 42322: contig of 42322 bp in length
42323 42422: gap of 100 bp
42423 63418: contig of 20996 bp in length
63419 63518: gap of 100 bp
63519 77969: contig of 14451 bp in length
77970 78069: gap of 100 bp
78070 93182: contig of 15113 bp in length
93183 93282: gap of 100 bp
93283 104093: contig of 10811 bp in length
104094 104193: gap of 100 bp
104194 115702: contig of 11509 bp in length
115703 115802: gap of 100 bp
115803 125183: contig of 9381 bp in length
125184 125283: gap of 100 bp
125284 132558: contig of 7275 bp in length
132559 137219: gap of 100 bp
137220 137319: contig of 4561 bp in length
137320 143276: gap of 100 bp
143277 143376: gap of 100 bp
143377 148922: contig of 5546 bp in length
148923 149022: gap of 100 bp
149023 154159: contig of 5137 bp in length
154160 154259: gap of 100 bp
154260 157686: contig of 3427 bp in length
157687 157786: gap of 100 bp
157787 159234: contig of 1448 bp in length
159235 159334: gap of 100 bp
159335 161039: contig of 1705 bp in length.

FEATURES

SOURCE

1. 161039
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-71066"
1. 42322
/note="assembly_fragment"
42423. 63418
/note="assembly_fragment"
63519. 77969
/note="assembly_fragment"
78070. 93182
/note="assembly_fragment"
93283. 104093
/note="assembly_fragment"
104194. 115702
/note="assembly_fragment"
115803. 125183
/note="assembly_fragment"
125284. 132558
/note="assembly_fragment"
132659. 137219
/note="assembly_fragment"
137320. 143276
/note="assembly_fragment"
143377. 148922
/note="assembly_fragment clone_end:77 vector_side:left"
149023. 154159
/note="assembly_fragment"
154260. 157686
/note="assembly_fragment"
157787. 159234
/note="assembly_fragment"
159335. 161039
/note="assembly_fragment"

BASE COUNT 49155 a 31166 c 31523 g 47795 t 1400 others
ORIGIN

Query Match 29.4%; Score 132.4; DB 2; Length 161039;
 Best Local Similarity 95.8%; Pred. No. 1.2e-19;
 Matches 136; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 148 ttagggaactacagatcctgtttggaaltatgacctttcttttggagttatcttcctt 207
 |||
 Db 42956 TTCAGACTATCCAGATCCGTTGGAATATGACCTTTCTTTGAGTATCTTCCTT 43015

QY 208 ttcaacttgttaaacacatatacaggttcccttataatttcttcaagatatccattc 267
 |||
 Db 43016 TTCACCTGTGTAACCAATATCCAGGTTCCCTTATATTCTTTCAGGATATCCATTC 43075

QY 268 ttgggctctgtttgttcat 289
 |||
 Db 43076 TGGGCTCTGTTTGGTGAGTA 43097

RESULT 15

ACUS 1179783

DEFINITION AX179783 246 bp DNA Sequence 3 from Patent WO0146417.

ACCESSION AX179783

VERSION AX179783.1 GI:15132145

KEYWORDS

SOURCE

ORGANISM

human.
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 246)

Walter, D.W. and Turner, C.A.

Human membrane proteins and polynucleotides encoding the same

patent: WO 0146417-A 3 28-JUN-2001;

Lexicon Genetics Incorporated (US)

Location/Qualifiers

1..246

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 60 a 45 c 47 g 94 t

ORIGIN

Query Match 25.1%; Score 113; DB 6; Length 246;

Best Local Similarity 100.0%; Pred. No. 2e-15;

Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 338 ttgggaatttgatatacatgacttgcacatatatgaattatcattctctgcctt 397
 |||
 Db 134 TGGGAATTTGATATGATGATGACTTTCAGCATTTATGATATTCATTCTCTGCTT 193

QY 398 ttctaatlttgggggtgcacatcagagattgtgattgtgaacaatgttgtga 450
 |||
 Db 194 TCTCAATTTTGGGTGTCACATCAGAGATTTGATGTGAACAATGTGTGTA 246

Search completed: November 28, 2001, 23:48:11
 Job time: 8134 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2001, 23:12:03 ; Search time 204.6 Seconds

(without alignments)
1885.613 Million cell updates/sec

Title: US-09-735-712-7

Perfect score: 450

Sequence: 1 atgattcaagcaccgcgcaca.....attgtgacaattgtgtga 450

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapept 1.0

Searched: 930621 seqs, 428662619 residues

tal number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	450	100.0	450	22	AA504282
2	347.4	77.2	603	22	AA504279
3	347.4	77.2	676	22	AA504283
4	347.4	77.2	747	22	AAH64741
5	340.8	75.7	417	22	AA504281
6	280.8	62.4	372	21	AA503080
7	234	52.0	468	22	AAH64745
8	113	25.1	246	22	AA504280
9	79	17.6	1563	21	AAZ36238
10	73.6	16.4	936	22	AAF58252
11	73.6	16.4	936	22	AAF58254

12	73.6	16.4	936	22	AAF58257	Oligonucleotide D1
13	73.6	16.4	936	22	AAF58259	Oligonucleotide D2
14	73.6	16.4	936	22	AAF58262	Oligonucleotide D2
15	73.6	16.4	936	22	AAF58255	Oligonucleotide D1
16	71.4	15.9	936	22	AAF58252	Oligonucleotide D1
17	71.4	15.9	936	22	AAF58254	Oligonucleotide D1
18	71.4	15.9	936	22	AAF58257	Oligonucleotide D1
19	71.4	15.9	936	22	AAF58259	Oligonucleotide D1
20	71.4	15.9	936	22	AAF58262	Oligonucleotide D2
21	71.4	15.9	936	22	AAF58255	Oligonucleotide D1
22	66.8	14.8	1330	21	AAH16693	Human secreted pro
23	66.8	14.8	1669	20	AAZ15104	Human high affinity
24	66.8	14.8	1670	21	AAZ32842	Human high affinity
25	50.2	11.2	708	22	AAF77694	Murine wild-type F
26	48.6	10.8	220	21	AAH45410	Human secreted exp
27	46	10.2	1661	17	AAH45120	Htm4 protein codin
28	46	10.2	1661	19	AAV05875	Htm4 gene. Homo s
29	44.4	9.9	244	22	AAF58238	Oligonucleotide D1
30	44.4	9.9	2545	11	AAO04645	Encodes beta subun
31	44.4	9.9	2545	12	AAO14734	Encodes beta subun
32	44.4	9.9	2545	14	AAO51021	Human FcERI beta g
33	43.6	9.7	244	22	AAF58238	Oligonucleotide D1
34	43	9.6	6888	21	AAH70114	Plasmodium falcipa
35	42	9.3	1474	13	AAQ21170	Clone CD20.4 encod
36	40.6	9.0	1060	20	AAH97730	Extended human sec
37	40.4	9.0	1473	10	AAH90610	Human CD20.4 anti
38	40.4	9.0	1474	17	AAH14710	Human cell surface
39	40.4	9.0	1474	21	AAH50584	Human lymphocyte c
40	40.4	9.0	1474	22	AAH503178	Human CD20 antigen
41	40.4	9.0	1476	10	AAH90613	Human CD20 antigen
42	40.4	9.0	1476	17	AAQ21173	Human CD20 antigen
43	40.4	9.0	1476	17	AAH14713	Human CD20 antigen
44	40.4	9.0	1476	19	AAH63451	Human CD20 antigen
45	40.4	9.0	1476	20	AAH81209	Human CD20 antigen

ALIGNMENTS

RESULT 1
ID AAS04282 standard; cDNA; 450 BP.
AC AAS04282;
XX 26-SEP-2001 (first entry)
XX
XX Novel human membrane protein #4 cDNA sequence.
DE Human; membrane protein; membrane receptor; Ige receptor; CD20;
XX Human; physiological disorder; ss.
XX
XX Homo sapiens.
OS
XX
XX
XX Key \Location/Qualifiers
FT 1.450 /tag= a
FT CDS /product= "Human membrane protein #4"
XX
XX WO200146417-A2.
XX
XX 28-JUN-2001.
XX
XX 12-DEC-2000; 2000MO-US33742.
XX
XX 22-DEC-1999; 99US-0171567.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Walke DW, Turner CA.
XX
XX WPI; 2001-408646/43.
XX
XX P-PSDB; AAU01213.

QY 301 ttccatattgcagtgaaagaaacacacagaaactctggaaatttgatcattgatg 360
 |||||
 DB 301 ttccatattgcagtgaaagaaacacacagaaactctgatattatgtgcgcgaataag 360
 QY 361 acttt 365
 ||||
 DB 361 aatct 365

RESULT 3

AA504283
 ID AA504283 standard; DNA; 676 BP.

AC AA504283;

DT 26-SEP-2001 (first entry)

DE DNA sequence encoding novel human membrane protein.

XX Human; membrane protein; membrane receptor; Ige receptor; CD20;
 XX physiological disorder; ds.

OS Homo sapiens.

PN WO200146417-A2.

PD 28-JUN-2001.

PF 12-DEC-2000; 2000WO-US33742.

PR 22-DEC-1999; 99US-0171567.

PA (LEXI-) LEXICON GENETICS INC.

PI Walke DW, Turner CA;

PI WPI; 2001-408646/43.

XX Polynucleotide encoding novel human membrane protein, useful for
 PT identifying agonist, antagonist or modifiers or for producing
 PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
 PT applications

PS Disclosure; Page 31-32; 32pp; English.

XX The present sequence encoding for a novel human membrane protein
 CC includes the flanking 5'- and 3'- sequences. Four amino acid sequences
 CC for novel human membrane proteins (AAU01210-AAU01213) are given in the
 CC present invention. These membrane proteins share structural similarity
 CC with membrane receptors such as the IGE receptor and mammalian CD20.
 CC The novel human membrane proteins are useful for identifying agonist,
 CC antagonist and modulators of the membrane proteins, and for producing
 CC antibodies specific to the membrane proteins. The membrane proteins can
 CC be used for diagnosis, drug screening, pharmacogenomic applications,
 CC clinical trial monitoring and the treatment of physiological disorders
 CC and diseases. The polynucleotides encoding the membrane proteins can be
 CC used to generate PCR primers or probes to identify mutations associated
 CC with a particular disease.

XX Sequence 676 BP; 193 A; 140 C; 112 G; 231 T; 0 other;

Query Match 77.2%; Score 347.4; DB 22; Length 676;

Best Local Similarity 97.0%; Pred. No. 2,9e-83;

Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 atgattcaagacgcacacagctcgcgtgttctgtatcttcctccaagaatcactgt 60
 |||||
 DB 33 atgattcaagacgcacacagctcgcgtgttctgtatcttcctccaagaatcactgt 92
 QY 61 tcaagatatgatccacagaaacttcacgcacagaccccttcaactcaagaagcccttgcaa 120
 |||||
 DB 93 tcaagatatgatccacagaaacttcacgcacagaccccttcaactcaagaagcccttgcaa 152

QY 121 aaatratctgtagaaaataatgaaatcttagggacratccagatccctgtttggaattag 180
 |||||
 DB 153 aaatratctgtagaaaataatgaaatcttagggacratccagatccctgtttggaattag 212
 QY 181 acccttctcttgaggtatctctcctttcaccctgtttaaacatatccaaagttccc 240
 |||||
 DB 213 acccttctcttgaggtatctctcctttcaccctgtttaaacatatccaaagttccc 272
 QY 241 ttatatctcttcagagatatccatctcgggctcgtttgttcaattaaatttgagcc 300
 |||||
 DB 273 ttatatctcttcagagatatccatctcgggctcgtttgttcaattaaatttgagcc 332
 QY 301 ttccatattgcagtgaaagaaacacacagaaactctggaaatttgatcattgatg 360
 |||||
 DB 333 ttccatattgcagtgaaagaaacacacagaaactctgatattatgtgcgcgaataag 392
 QY 361 acttt 365
 ||||
 DB 393 aatct 397

RESULT 4

AAH64741
 ID AAH64741 standard; CDNA; 747 BP.

AC AAH64741;

DT 11-SEP-2001 (first entry)

DE Human secreted protein CDNA, SEQ ID NO: 17.

XX Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
 XX GENSET; ss.

XX Homo sapiens.

PN WO200142451-A2.

PD 14-JUN-2001.

PF 07-DEC-2000; 2000WO-IB01938.

PR 08-DEC-1999; 99US-0169629.

PR 06-MAR-2000; 2000US-0187470.

PA (GENSET) GENSET.

PI Dumas Milne Edwards J, Bougueleret L, Jobert S;

PI WPI; 2001-367870/38.

PI P-PSDB; AA689138.

XX Full length GENSET human nucleic acids encoding potentially secreted
 PT proteins, useful in gene therapy and vaccination against a variety of
 PT diseases, and for diagnosis of those diseases -

PS Claim 7; Page 576-577; 921pp; English.

XX The invention relates to full length GENSET human nucleic acids encoding
 CC potentially secreted proteins. The nucleic acids and the polypeptides
 CC they encode may be used in the prevention, treatment and diagnosis of
 CC diseases associated with inappropriate GENSET gene expression. For
 CC example, they be used to treat disorders associated with decreased
 CC GENSET gene expression by rectifying mutations or deletions in a
 CC patient's genome that affect the activity of GENSET or by supplementing
 CC the patient's own production of GENSET polypeptides. Conversely,
 CC antisense nucleic acid molecules may be administered to down regulate
 CC GENSET expression by binding with the cells' own genes and preventing
 CC their expression. The sense and antisense nucleic acids may also be
 CC used as DNA probes in diagnostic assays to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples, and hence to
 CC determine which patients may be in need of restorative therapy.

CC The GENSET polypeptides may be used as antigens in the production of
 CC antibodies and in assays to identify modulators (agonists and
 CC antagonists) of GENSET polypeptide expression and activity. The
 CC present sequence is a GENSET nucleic acid of the invention.
 CC
 SQ Sequence 747 BP; 240 A; 145 C; 121 G; 241 T; 0 other;

Query Match 77.2%; Score 347.4; DB 22; Length 747;
 Best Local Similarity 97.0%; Pred. No. 3e-83;
 Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1 atgagatcaagcaccgacacagtcggtgttctgtgtatctccccaagaatcactgct 60
 |||
 Db 73 atgagatcaagcaccgacacagtcggtgttctgtgtatctccccaagaatcactgct 132
 OY 61 tcagaataatgagtcacagaaacttcagcagacacttcacccaagaagcccttgcaa 120
 |||
 Db 133 tcagaataatgagtcacagaaacttcagcagacacttcacccaagaagcccttgcaa 192
 OY 121 aaatattgtctagaagaataaactcttaggactatccagactcctgttggaattatg 180
 |||
 Db 193 aaatattgtctagaagaataaactcttaggactatccagactcctgttggaattatg 252
 OY 181 acccttcttttgaggtatctctcttccacttgtaaaacacatccaggtttccc 240
 |||
 Db 253 acccttcttttgaggtatctctcttccacttgtaaaacacatccaggtttccc 312
 OY 241 ttatattcttccagagatccatctctggtctgtttgttcaatcaatctgagacc 300
 |||
 Db 313 ttatattcttccagagatccatctctggtctgtttgttcaatcaatctgagacc 372
 OY 301 ttcccaattgcagtgaaaagaacacacagaaactctgggaatttgaattcatgtatg 360
 |||
 Db 373 ttcccaattgcagtgaaaagaacacacagaaactctgggaatttgaattcatgtatg 432
 OY 361 acttt 365
 |||
 Db 433 aattt 437

RESULT 5
 ID AAS04281 standard; cDNA; 417 BP.
 AC AAS04281;

26-SEP-2001 (first entry)
 Novel human membrane protein #3 cDNA sequence.

Human: membrane protein; membrane receptor; IGF receptor; CD20;
 KW physiological disorder; ss.

OS Homo sapiens.

Key Location/Qualifiers
 FT 1..417
 FT CDS /*tag= a
 FT /product= "Human membrane protein #3"

W0200146417-A2.

28-JUN-2001.

12-DEC-2000; 2000WO-US33742.

22-DEC-1999; 99US-0171567.

(LEXI-) LEXICON GENETICS INC.

Walke DW, Turner CA;

DR WPI: 2001-408646/43.
 DR P-PSDB: AA001212.
 XX
 XX Polynucleotide encoding novel human membrane protein, useful for
 PT identifying agonist, antagonist or modifiers or for producing
 PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
 PT applications
 XX
 XX Disclosure: Page 30; 32pp; English.

The present sequence encodes for novel human membrane protein #3.
 Human membrane protein #3 is 1 of 4 human membrane proteins
 (AA001210-AA001213) given in the present invention. These membrane
 proteins share structural similarity with membrane receptors such as
 the IGF receptor and mammalian CD20. The novel human membrane proteins
 are useful for identifying agonists, antagonists and modulators of the
 CC membrane proteins, and for producing antibodies specific to the
 CC membrane proteins. The membrane proteins can be used for diagnosis,
 CC drug screening, pharmacogenomic applications, clinical trial monitoring
 CC and the treatment of physiological disorders and diseases. The
 CC polynucleotides encoding the membrane proteins can be used to generate
 CC PCR primers or probes to identify mutations associated with a particular
 CC disease.

SQ Sequence 417 BP; 115 A; 91 C; 69 G; 142 T; 0 other;

Query Match 75.7%; Score 340.8; DB 22; Length 417;
 Best Local Similarity 94.1%; Pred. No. 1.5e-81;
 Matches 354; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

OY 1 atgagatcaagcaccgacacagtcggtgttctgtgtatctccccaagaatcactgct 60
 |||
 Db 1 atgagatcaagcaccgacacagtcggtgttctgtgtatctccccaagaatcactgct 60
 OY 61 tcagaataatgagtcacagaaacttcagcagacacttcacccaagaagcccttgcaa 120
 |||
 Db 61 tcagaataatgagtcacagaaacttcagcagacacttcacccaagaagcccttgcaa 120
 OY 121 aaatattgtctagaagaataaactcttaggactatccagactcctgttggaattatg 180
 |||
 Db 121 aaatattgtctagaagaataaactcttaggactatccagactcctgttggaattatg 180
 OY 181 acccttcttttgaggtatctctcttccacttgtaaaacacatccaggtttccc 240
 |||
 Db 181 acccttcttttgaggtatctctcttccacttgtaaaacacatccaggtttccc 240
 OY 241 ttatattcttccagagatccatctctggtctgtttgttcaatcaatctgagacc 300
 |||
 Db 241 ttatattcttccagagatccatctctggtctgtttgttcaatcaatctgagacc 300
 OY 301 ttcccaattgcagtgaaaagaacacacagaaactctgggaatttgaattcatgtatg 360
 |||
 Db 301 ttcccaattgcagtgaaaagaacacacagaaactctgggaatttgaattcatgtatg 360
 OY 361 acttcagcattatg 376
 |||
 Db 361 attccacaaaatag 376

RESULT 6
 ID AAC03080 standard; cDNA; 372 BP.
 AC AAC03080;

06-OCT-2000 (first entry)

Human secreted protein 5' EST, SEQ ID NO: 3078.

Human: 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.

OS Homo sapiens.
XX
XX EPI033401-A2.
PN
XX 06-SEP-2000.
PD
XX
XX 21-FEB-2000; 2000EP-0200610.
PF
XX
XX 26-FEB-1999; 99US-0122487.
PR
XX
XX (GENSET).
PA
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI
XX
XX WPI: 2000-500381/45.
DR
XX P-PSDB; AAG03074.
XX
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
PT
XX
XX Claim 1: SEQ ID 3078; 71pp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from
CC cDNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
CC
XX
XX

Sequence 372 BP; 105 A; 87 C; 53 G; 117 T; 10 other;

Query Match 62.4%; Score 280.8; DB 21; Length 372;
Best Local Similarity 96.3%; Pred. No. 1.3e-65;
Matches 289; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 1 atggattcaagaccgacacagtcggtgttcgtgatttcctccgaataactgtc 60
DB 73 atggattcaagaccgacacagtcggtgttcgtgatttcctccgaataactgtc 132
DB 61 tcagatatgagtcacagaaacttcagacagacatttcacccaagcccttgca 119
DB 133 tcaraatatgagtcacagaaacttcagacagacatttcacccaagcccttgca 192
QY 120 aaatatattgtcagaaaatgaataactttaggaactaccagatcctgttgaattat 179
DB 193 aaatatattgtcagaaaatgaataactttaggaactaccagatcctgttgaattat 252
QY 180 gacccttctcttggaagtattctcttcccttgaactgttaaacatatccaaagtttc 239
DB 253 gacccttctcttggaagtattctcttcccttgaactgttttgaactgttgaattat 312
QY 240 cttatatctcttcagatatccattctgggctcgtgttgcattcaattcggagc 299
DB 313 cttatatctcttcagatatccattctgggctcgtgttgcattcaattcggagc 372

RESULT 7
AAH64745
ID AAH64745 standard; cDNA; 468 BP.
XX
XX AAH64745;
XX
XX 11-SEP-2001 (first entry)

XX Human secreted protein cDNA, SEQ ID NO: 21.
DE
XX
XX Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KM GENSET; ss.
KM
XX
XX Homo sapiens.
OS
XX
XX WO200142451-A2.
PN
XX
XX 14-JUN-2001.
PD
XX
XX 07-DEC-2000; 2000MO-1B01938.
PF
XX
XX 08-DEC-1999; 99US-0169629.
PR
XX 06-MAR-2000; 2000US-0187470.
XX
XX (GENSET).
PA
XX
XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
PI
XX
XX WPI: 2001-367870/38.
DR
XX P-PSDB; AAG89142.
XX

Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases -
PT
XX
XX
XX Claim 7: Page 581; 921pp; English.

XX The invention relates to full length GENSET human nucleic acids encoding
XX potentially secreted proteins. The nucleic acids and the polypeptides
XX they encode may be used in the prevention, treatment and diagnosis of
XX diseases associated with inappropriate GENSET gene expression. For
XX example, they be used to treat disorders associated with decreased
XX GENSET gene expression by rectifying mutations or deletions in a
XX patient's genome that affect the activity of GENSET or by supplementing
XX the patient's own production of GENSET polypeptides. Conversely,
XX antisense nucleic acid molecules may be administered to down regulate
XX GENSET expression by binding with the cells' own genes and preventing
XX their expression. The sense and antisense nucleic acids may also be
XX used as DNA probes in diagnostic assays to detect and quantitate the
XX presence of similar nucleic acid sequences in samples, and hence to
XX determine which patients may be in need of restorative therapy.
XX The GENSET polypeptides may be used as antigens in the production of
XX antibodies and in assays to identify modulators (agonists and
XX antagonists) of GENSET polypeptide expression and activity. The
XX present sequence is a GENSET nucleic acid of the invention.
XX
XX

Sequence 468 BP; 155 A; 94 C; 81 G; 138 T; 0 other;

Query Match 52.0%; Score 234; DB 22; Length 468;
Best Local Similarity 77.1%; Pred. No. 4.1e-53;
Matches 347; Conservative 0; Mismatches 0; Indels 103; Gaps 1;

QY 1 atggattcaagaccgacacagtcggtgttcgtgatttcctccgaataactgtc 60
DB 76 atggattcaagaccgacacagtcggtgttcgtgatttcctccgaataactgtc 135
QY 61 tcagatatgagtcacagaaacttcagacagacatttcacccaagcccttgca 120
DB 136 tcagatatgagtcacagaaacttcagacagacatttcacccaagcccttgca 195
QY 121 aaatatattgtcagaaaatgaataactttaggaactaccagatcctgttgaattat 180
DB 196 aaatatattgtcagaaaatgaataactttaggaactaccagatcctgttgaattat 227
QY 181 acccttctcttggaagtattctcttcccttgaactgttgaataaacatatccaaagtttc 240
DB 228 ----- 227
QY 241 ttatatattcttcagatatccattctgggctcgtgttgcattcaattcggagc 300

Db 228 -----ggatatacattctgggctctgttctgttctcaattatcttgagcc 272
QY 301 ttccctaatgcaagtgaagaaacacacagaacatcttgggaatttgcattatcatgacg 360.
Db 273 ttccctaatgcaagtgaagaaacacacagaacatcttgggaatttgcattatcatgacg 332
QY 361 acttcaagcattatgaattatctctgccttctcaatttggggcgaccta 420
Db 333 acttcaagcattatgaattatctctgccttctcaatttggggcgaccta 392
QY 421 gaggattgtgattgtgaacaatgtgttga 450
Db 393 gaggattgtgattgtgaacaatgtgttga 422

RESULT 8

AAS04280
ID AAS04280 standard; cDNA; 246 BP.

AAS04280;

26-SEP-2001 (first entry)

Novel human membrane protein #2 cDNA sequence.

Human; membrane protein; membrane receptor; IgE receptor; CD20;
physiological disorder; ss.

Homo sapiens.

Key Location/Qualifiers

1..246

/*tag= a

/product= "Human membrane protein #2"

W0200146417-A2.

28-JUN-2001.

12-DEC-2000; 2000WO-US33742.

22-DEC-1999; 99US-0171567.

(LEXI-) LEXICON GENETICS INC.

Walke DW, Turner CA;

WPI; 2001-408646/43.

P-PSDB; AAU01211.

Polynucleotide encoding novel human membrane protein, useful for
identifying agonist, antagonist or modifiers or for producing
antibodies useful in therapeutic, diagnostic and pharmacogenomic
applications

Disclosure; Page 30; 32pp; English.

The present sequence encodes for novel human membrane protein #2.

Human membrane protein #2 is 1 of 4 human membrane proteins
(AAU01210-AAU01213) given in the present invention. These membrane
proteins share structural similarity with membrane receptors such as
the IgE receptor and mammalian CD20. The novel human membrane proteins
are useful for identifying agonists, antagonists and modulators of the
membrane proteins, and for producing antibodies specific to the
drug screening, pharmacogenomic applications, clinical trial monitoring
and the treatment of physiological disorders and diseases. The
polynucleotides encoding the membrane proteins can be used to generate
PCR primers or probes to identify mutations associated with a particular
disease.

Sequence 246 BP; 60 A; 45 C; 47 G; 94 T; 0 other;

Query Match 25.1%; Score 113; DB 22; Length 246;
Best Local Similarity 100.0%; Pred. No. 5.6e-21;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 338 tgggaatttgcattatcatgagacttcaagcattatgaattatcatctctgcctt 397
Db 134 tgggaatttgcattatcatgagacttcaagcattatgaattatcatctctgcctt 193

QY 398 tctcaatttggggtgcacactcagaagattgtgattgtgaacaatgtgttga 450
Db 194 tctcaatttggggtgcacactcagaagattgtgattgtgaacaatgtgttga 246

RESULT 9

AAZ36238
ID AAZ36238 standard; cDNA; 1563 BP.

AAZ36238;

22-FEB-2000 (first entry)

cDNA encoding a bone marrow secreted protein designated BMS208.

Bone marrow secreted protein; bone marrow stromal cell; cytokine;
cell proliferation; cell differentiation; hematopoiesis; anaemia;
myeloid cell deficiency; lymphoid cell deficiency; myeloid cell;
erythroid progenitor cell; colony stimulating factor; granulocyte;
monocyte; macrophage; myelo-suppression; megakaryocyte; platelet;
platelet disorder; thrombocytopenia; hematopoietic stem cell;
stem cell disorder; aplastic anaemia; bone differentiation;
paroxysmal nocturnal hemoglobinuria; bone growth; cartilage; tendon;
ligament; nerve; wound healing; tissue repair; burn; incision; ulcer;
bone fracture; cartilage damage; artificial joint; ss.

Homo sapiens.

Key Location/Qualifiers

255..758

/*tag= a

/product= "bone marrow secreted protein"

sig_peptide 255..440

/*tag= b

polya_signal 1531..1536

/*tag= c

W09933979-A2.

08-JUL-1999.

18-DEC-1998; 98WO-US27008.

30-DEC-1997; 97US-0068958.

PR 24-SEP-1998; 98US-0101603.

PR 30-SEP-1998; 98US-0102540.

(CHIR) CHIRON CORP.

Lin H, Cao L;

WPI; 2000-038344/03.

P-PSDB; AAY3632.

New isolated human polynucleotide and secreted proteins can induce
production of other cytokines in certain cell populations

Claim 11; Page 92-94; 120pp; English.
AAZ36228-49 encode bone marrow secreted proteins of human bone marrow
stromal cells. The proteins can exhibit cytokine, cell proliferation, or
cell differentiation activity (either inducing or inhibiting). They can
be used to support colony forming cells or factor-dependent cell lines,

24-APR-2001 (first entry)
Oligonucleotide D2004.
Electron-transfer group; ETW; mismatch; genotyping;
gene expression; ss.
Synthetic.
WO200107665-A2.
01-FEB-2001.
26-JUL-2000; 2000WO-US20476.
26-JUL-1999; 99US-0145695.
17-MAR-2000; 2000US-0190259.
(CLIN-) CLINICAL MICRO SENSORS INC.
Umek RM;
WPI; 2001-159728/16.
Nucleic acids containing electron-transfer group, useful as labels in
hybridization assays, e.g. for genotyping, allowing repeat analyses on
a single surface -
Example 6; Page 128; 159pp; English.
The present invention relates to a composition comprising two nucleic
acids each containing an electron-transfer group (ETW) having
different redox potentials. The invention is used for electronic
detection of nucleic acids, especially of substitutions (mismatches)
and single-nucleotide polymorphisms, e.g. for genotyping,
monitoring gene expression.
Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;

[illegible]

RESULT 14
 AAF58262
 ID AAF58262 standard; DNA; 936 BP.
 XX
 AC AAF58262;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Oligonucleotide D2007.
 XX
 KW Electron-transfer group; ETM; mismatch; genotyping;
 KW gene expression; ss.
 XX
 OS Synthetic.
 XX
 PN W0200107665-A2.
 XX
 PD 01-FEB-2001.
 XX
 PF 26-JUL-2000; 2000WO-US20476.
 XX
 PR 26-JUL-1999; 9905-0145695.
 PR 17-MAR-2000; 2000US-0190259.
 XX
 PA (CLIN-) CLINICAL MICRO SENSORS INC.
 XX
 PI Umek RM;
 XX
 DR WPI: 2001-159728/16.
 XX
 PT Nucleic acids containing electron-transfer group, useful as labels in
 PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
 PT a single surface -
 XX
 PS Example 6; Page 128; 159pp; English.
 XX
 CC The present invention relates to a composition comprising two nucleic
 CC acids each containing an electron-transfer group (ETM) having
 CC different redox potentials. The invention is used for electronic
 CC detection of nucleic acids, especially of substitutions (mismatches)
 CC and single-nucleotide polymorphisms, e.g. for genotyping,
 CC monitoring gene expression.
 XX
 SQ Sequence 936 BP; 5 A; 139 C; 10 G; 6 T; 776 other;

Query Match Similarity 16.4%: Score 73.6: DB 22: Length 936:
 Best Local Similarity 1.7%: Pred No. 2,6e-10:
 Matches 7: Conservative 261: Mismatches 150: Indels 0: Gaps 0:

```

QY 31 ttctcgtatctcctccgaacacgtcgttcagatatgagtcacagaaattcagcc 90
   :: :::: : : :::: : : :::: : : :::: : : :::: : : ::::
Db 75 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 134

QY 91 acgaccttcaacccaagcccttgaaaaatatctgcagaaaaatgaaattcta 150
   : : :::: : : ::: : : ::::: : : ::::: : : ::::: : : ::::
Db 135 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 194

QY 151 gggactacacagatcctgttgaaatagaccttctttcgtgagttactcttc 210
   ! : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 195 wgvwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 254

QY 211 acctgttaaaaccatccaagttcccttatttcttcagatcatcattcgg 270
   : : ::::: : : : : : : : : : : : : : : : : : : : : :
Db 255 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 314

QY 271 ggcctgtttgttactaatctcgagccttcctcaatgcagtgaaaaaaccca 330
   : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 315 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 374

QY 331 gaacctcgggaatttgattacatgatgacctcagcatatatgaattatctct 390
   : : : : : : : : : : : : : : : : : : : : : : : : : :
  
```

Db 375 wwwgcttcaatcctcgaagcctcctcaatcgaagtgagaaagaaaccaca 434
 Qy 331 ctgcttctcgaatttgagggtgcacacgaaggtgtgaltggaacatgtgt 448
 Db 435 wwwgcttcaatcctcgaagcctcctcaatcgaagtgagaaagaaaccaca 492

RESULT 15

AAFS8255 ID AAF58255 standard; DNA; 938 BP.

AAF58255; AAF58255;

24-APR-2001 (first entry)

XX XX Oligonucleotide D1876.

XX XX Electron-transfer group; ETM; mismatch; genotyping; gene expression; ss.

XX XX Synthetic.

XX XX W0200107665-A2.

XX XX 01-FEB-2001.

XX XX 26-JUL-2000; 2000MO-US20476.

XX XX 26-JUL-1999; 99US-0145695.

XX XX 17-MAR-2000; 2000US-0190259.

XX XX (CLIN-) CLINICAL MICRO SENSORS INC.

XX XX Umek RM;

XX XX WPI; 2001-159728/16.

XX XX Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface

XX XX Example 6; Page 127; 159p; English.

XX XX The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.

SO Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;

Query Match 16.4%; Score 73.6; DB 22; Length 938;
 Best Local Similarity 1.7%; Pred. No. 2.6e-10;
 Matches 7; Conservative 261; Mismatches 150; Indels 0; Gaps 0;

Qy 31 ttcttgatctcctcgaatcactgctcgaataatgagtcacagaacttcagcc 90
 Db 75 wwwgcttcaatcctcgaagcctcctcaatcgaagtgagaaagaaaccaca 134
 Qy 91 acgaacttcaactcaagcccttgcaaaaattgtgtagaaaatgaactcta 150
 Db 135 wwwgcttcaatcctcgaagcctcctcaatcgaagtgagaaagaaaccaca 194
 Qy 151 gggactacagatccttgtagaatatgacacttctcttgagtgatcctcttc 210
 Db 195 wwwgcttcaatcctcgaagcctcctcaatcgaagtgagaaagaaaccaca 254
 Qy 211 acctgttaaacacatcgaaggttccttatattcttcagatatacatctcg 270
 Db 255 wwwgcttcaatcctcgaagcctcctcaatcgaagtgagaaagaaaccaca 314

Qy 271 ggcctgttctcgaatcctcgaagcctcctcaatcgaagtgagaaagaaaccaca 330
 Db 315 wwwgcttcaatcctcgaagcctcctcaatcgaagtgagaaagaaaccaca 374
 Qy 331 gaactctgggaatttgatatactgagacttcagcatatggaattatctct 390
 Db 375 wwwgcttcaatcctcgaagcctcctcaatcgaagtgagaaagaaaccaca 434
 Qy 391 ctgcttctcgaatcctcgaagcctcctcaatcgaagtgagaaagaaaccaca 448
 Db 435 wwwgcttcaatcctcgaagcctcctcaatcgaagtgagaaagaaaccaca 492

Search completed: November 29, 2001, 00:23:54
 Job time: 4311 sec

THIS PAGE BLANK (USPTO)

Query Match 10.2%; Score 46; DB 1; Length 1661;
Best Local Similarity 55.2%; Pred. No. 0.00074;
Matches 116; Conservative 0; Mismatches 85; Indels 9; Gaps 1;

QY 135 aaaaatgaataatcttaggaactacagatccctgttggaattatgacctttctttg 194
DB 234 AAAATTACAAAGTCTTGCGGCCATCCAGATCCGATGACGACATGATTCTGGCTTTGGG 293
QY 195 agt-----tatcttcctttacacctgtttaaacatatccaagttcccttat 245
DB 294 TGCTTTCTGGGTTCCCTTCAATACCCATACCACTTCCAAAAGCACTCTTTTCTTCAC 353
QY 246 attcttcaggaatcatctctgggacctgtttgttcatatattctcgagacctct 305
DB 354 CTCTACACAGGCTACCCGATTTGGGGTGTGTTTCTGTAGTTCAGGAACCTTGTTC 413
QY 306 aattgcagtgtaaaagaaaacacagagaac 335
DB 414 TGTTGTACGAGGATTAACCCACAGAGAAC 443

RESULT 4

US-08-707-340-3
Sequence 3, Application US/08707340
Patent No. 5705615

GENERAL INFORMATION:

APPLICANT: Lim, Bing
APPLICANT: Adra, Chaker N.
APPLICANT: Lelias, Jean-Michel
TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
TITLE OF INVENTION: ASSAYS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/707,340
FILING DATE: 03-SEP-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,492
FILING DATE: 06-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/675,648
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BH94-03A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 97..738
US-08-707-340-3

Query Match 10.2%; Score 46; DB 1; Length 1661;
Best Local Similarity 55.2%; Pred. No. 0.00074;
Matches 116; Conservative 0; Mismatches 85; Indels 9; Gaps 1;

QY 135 aaaaatgaataatcttaggaactacagatccctgttggaattatgacctttctttg 194
DB 234 AAAATTACAAAGTCTTGCGGCCATCCAGATCCGATGACGACATGATTCTGGCTTTGGG 293
QY 195 agt-----tatcttcctttacacctgtttaaacatatccaagttcccttat 245
DB 294 TGCTTTCTGGGTTCCCTTCAATACCCATACCACTTCCAAAAGCACTCTTTTCTTCAC 353
QY 246 attcttcaggaatcatctctgggacctgtttgttcatatattctcgagacctct 305
DB 354 CTCTACACAGGCTACCCGATTTGGGGTGTGTTTCTGTAGTTCAGGAACCTTGTTC 413
QY 306 aattgcagtgtaaaagaaaacacagagaac 335
DB 414 TGTTGTACGAGGATTAACCCACAGAGAAC 443

RESULT 5

US-08-994-578-3
Sequence 3, Application US/08994578
Patent No. 5972688

GENERAL INFORMATION:

APPLICANT: Lim, Bing
APPLICANT: Adra, Chaker N.
APPLICANT: Lelias, Jean-Michel
TITLE OF INVENTION: HTM4 METHODS OF TREATMENT AND ASSAYS,
TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/994,578
FILING DATE: December 19, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/707,340
FILING DATE: 03-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/675,648
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BH94-03A2Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:

Accession	Sequence	Position
Dd	cttcgaacttggagacgaatgctttattatataatgaagcagcgtacccttctgggtg	367
Oy	ctgttttgatcaatctctgagccttcctaattgcagtgaagaacacacagaaa	334
Dd	cagtcgtctgtttgtctgataatttgtcaattatgctcgaagaagaacacacatgt	427
Oy	ctctctg	340
Dd	atctgg	433

RESULT 8

US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHRIFFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9c-Fls
US-08-232-463-14

Query Match	9.2%;	Score 41.6;	DB 1;	Length 7218;
Best Local Similarity	4.2%;	Pred. No. 0.02;		
Matches 17;	Conservative 214;	Mismatches 173;	Indels 0;	Gaps 0;

[illegible]

Oy 126 attgcgcagaaaatgaaatccttaggcagcattccagatccctgtttgagatagacct 185
 Db 1152 YY 1211
 Oy 186 tctcttggagltactctcctttcacctgtttaaaccatalccagaagttcccttat 245
 Db 1212 YY 1271
 Oy 246 attccttcagatccattcttgaggctctgtttgttcataatctcgagcctct 305
 Db 1272 YY 1331
 Oy 306 aattgcagltgaaaagaaaaccacagaactctgggaatttgattatcatgatcctt 365
 Db 1332 YY 1391
 Oy 366 cagcatattgaattatcatctctcgtccttccaatttgg 409
 Db 1392 YYG 1435

RESULT 9

```

US-09-104-324B-1/C
: Sequence 1, Application US/09104324B
: Patent No. 6232460
:
: GENERAL INFORMATION:
: APPLICANT: T recel, Ozlem; Sahin, Ugur; Pfreundschuh, Michael
: TITLE OF INVENTION: Methods For Diagnosis And Treating Cancers,
: TITLE OF INVENTION: And Methods For Identifying Pathogenic Markers-In A Sample
: TITLE OF INVENTION:
: NUMBER OF SEQUENCES: 4
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fulbright & Jaworski LLP
: STREET: 666 Fifth Avenue
: CITY: New York City
: STATE: New York
: ZIP: 10103
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
:
: COMPUTER: IBM
: OPERATING SYSTEM: PC-DOS
: SOFTWARE: Wordperfect
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/104,324B
: FILING DATE: 25-June-1998
: CLASSIFICATION: 435
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/892,702
: FILING DATE: 15-July-1997
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Hanson, No. 6232460man D.
: REGISTRATION NUMBER: 30,946
: REFERENCE/DOCKET NUMBER: LUD 5491
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 318-3000
: TELEFAX: (212) 752-5598
:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3393 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
:
US-09-104-324B-1

```

Query Match	7.78;	Score 34.8;	DB 4;	length 3393;
Best Local Similarity	49.58;	Pred. No. 1.2;		
Matches 90;	Conservative 0;	Mismatches 92;	Indels 0;	Gaps 0;

QY 186 ttccttttgagttatccttccttttcaacctgtaaacacatcacaaggtttcccttat 245
||||| | | | | | | | | | | | | | | | | |
Db 3004 ttcttttagtttttttttttctatccattttttagcaaatatcagccccaacggctctccggcat 2945

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2001, 21:34:03 ; Search time 1970.99 Seconds
(without alignments)
2453.387 Million cell updates/sec

Title: US-09-735-712-7

Perfect score: 450

Sequence: 1 atgtatcaagcaccgcacac.....atgtgacaacatgtgtga 450

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 segs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estlom:*
5: em_estlpl:*
6: em_estlba:*
7: em_estlro:*
8: em_estlov:*
9: em_estc:*
10: qb_est1:*
11: qb_est2:*
12: qb_hlc:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	370	82.2	410	10	AA411806 zt67a03.s
C 2	348	77.3	389	10	AA781801 a158901.s
C 3	347.4	77.2	516	10	AA436088 z003a08.r
C 4	215.4	47.9	508	10	AT149899 qf43b06.x
C 5	214.8	47.7	482	10	AT1002083 ol38b02.s
C 6	197.2	43.8	281	10	AA707529 ah41a12.s
C 7	157.8	33.1	415	10	AA470059 zt94h05.r
C 8	145.8	32.4	484	10	BE107659 UT-R-BT1-
C 9	126.2	28.0	224	10	BE638325 EST00022
C 10	113	25.1	382	10	AA435988 z003a08.s
C 11	113	25.1	387	10	AA416972 zt94h05.s
C 12	112.4	25.0	538	13	AO108532 C17-HSP-2

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
AA411806	410 bp mRNA	AA411806	1	EST	human.	Human sapiens	1 (bases 1 to 410)	Hillier, L., Allen, M., Bowles, L., Dubuque, T., Getzel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.	Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.edu	High quality sequence stop: 401.	Location/Qualifiers 1. 410
AA411806	410 bp mRNA	AA411806	1	EST	human.	Human sapiens	1 (bases 1 to 410)	Hillier, L., Allen, M., Bowles, L., Dubuque, T., Getzel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.	Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.edu	High quality sequence stop: 401.	Location/Qualifiers 1. 410

ALIGNMENTS

AA411806 410 bp mRNA EST 12-AUG-1997
zt67a03.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:777372
3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;
mRNA sequence.
AA411806
AA411806.1 GI:2070377
EST.
human.
Human sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 410)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Getzel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 608 Std Error: 0.00
Seq primer: ~41m3 fwd. Et from Amersham
High quality sequence stop: 401.
Location/Qualifiers
1. 410
/organism="Homo sapiens"
/db_xref="GDB:5924284"

REFERENCE 1 (bases 1 to 516)
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, J., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepcoe, M., Tan, F., Theising, B., White, T., Wylie, T., Waterston, R. and Wilson, R.
 TITLE WashU-NCI human EST Project
 JOURNAL Unpublished (1997)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 733 Std Error: 0.00
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 482.

FEATURES
 source
 1..516
 /organism="Homo sapiens"
 /db_xref="GDB:5927437"
 /db_xref="taxon:9606"
 /clone="IMAGE:730742"
 /clone.lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: p773D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCGAATGAGGAGCGGCCGCCCAATTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaído."

BASE COUNT 149 a 117 c 80 g 170 t
 ORIGIN

Query Match 77.2%; Score 347.4; DB 10; Length 516;
 Best Local Similarity 97.0%; Pred. No. 4.1e-68;
 Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 atggatcaagaacgcagtcgcgtgttctgtgtatttcctccagaatacactgct 60
 |||
 Db 52 ATGGATTCAAGACCGCACAGTCGGTGTGTTGTTCTTCACAGAAATCACTGCT 111
 |||
 QY 61 tcagaatagatgcacagaacttcagcagcactttcaactcaagcccttgcaa 120
 |||
 Db 112 TCAGAAATAGATCCACAGAACTTTCAGCCAGACCTTTTCACACTCAAGCCCTTGCA 171
 |||
 QY 121 aaattatctctagaanaaaatctaggaactatccagatccctgttggaaattg 180
 |||
 Db 172 AAATATTGTGTAGAAATGAAATCTTAGGACTATCCAGATCCTGTTGGAATTAG 231
 |||
 QY 181 accttctcttggagttatctctcttccactgttaaaacatataccaaggttccc 240
 |||
 Db 232 ACCCTTCTTTGGAGTATCTTCTTCACTTGTGTTAAACCAATCAAGGTTTCCC 291
 |||
 QY 241 ttataatctcttggagatcatcttgggctctgttttcttaattatctggaacc 300
 |||
 Db 292 TTTATATTCTTTCAGAGATATCTTCTGAGGCTCTGTTTGTTCATTATTTGAGGCC 351
 |||
 QY 301 ttccctaattgcagtgaaaagaaaacacagaaactctggaaatttgattacattg 360
 |||
 Db 352 TTCCTAATTCAGTGAAGAAAACACAGAACTCTGATATATTGAGCCGAATTAAG 411
 |||
 QY 361 acttt 365
 |||
 Db 412 AATTT 416

RESULT 4
 LOCUS A1149899/c
 DEFINITION q14h06.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1752827
 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;
 mRNA sequence.
 A1149899
 A1149899.1 GI:3678368
 EST.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 508)
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaído
 , Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 www.bio.lnl.gov/btrp/image/image.html
 Insert Length: 742 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 455.

FEATURES
 source
 1..508
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1752827"
 /clone.lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: p773D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCGAATGAGGAGCGGCCGCCCAATTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaído."

BASE COUNT 186 a 91 c 89 g 142 t
 ORIGIN

Query Match 47.9%; Score 215.4; DB 10; Length 508;
 Best Local Similarity 95.3%; Pred. No. 1.6e-38;
 Matches 222; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 133 agaaaaatgaaaactctaggaactatccagatccctgttggaaattagacctttctt 192
 |||
 Db 508 AGAAAAATGAAATCTTAGGACTATCCAGATCCTTGGAAATATGACCTTCTT 449
 |||
 QY 193 ggaattatctcttcttccactgttaaaacatataccaaggttcccttatattct 252
 |||
 Db 448 GGAATATCTTCTTTCAGACTTGTGTTAAACCAATCAAGGTTTCCCTTATATTTCT 389
 |||
 QY 253 tcaggaatccattctggggtctctgttcttcaatattcttggagccttccaaattga 312
 |||
 Db 388 TCAGGATATCCATTCGAGGCTCTGTTTGTTCATTAATTGAGACCTTCTTAATTGCA 329
 |||
 QY 313 gtgaaaagaaaacacagaaactctgggaatttgattacattgacctt 365
 |||
 Db 328 GTGAAAAGAAAACACAGAACTCTGATATATTGAGCCGAATATATT 276

```

RESULT 5
A1002083/c 482 bp mRNA EST 27-AUG-1998
LOCUS 013b02.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1619019
DEFINITION 3', mRNA sequence.
ACCESSION A1002083
VERSION A1002083.1 GI:3202120
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 482)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TUMOR National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
FEATURES
SOURCE
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
Insert Length: 598 Std Error: 0.00
Seq primer: -40m13 fwd. EF from Amersham
High quality sequence stop: 393.
Location/Qualifiers
1..482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1619019"
/clone_id="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTTCACATCTGAGTGAAGGAGCGCGCCCAATTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 157 a 85 c 99 g 141 t
ORIGIN
Query Match 47.7%; Score 214.8; DB 10; Length 482;
Best Local Similarity 70.7%; Pred. No. 2.2e-38;
Matches 318; Conservative 0; Mismatches 97; Indels 35; Gaps 1;
QY 1 atgagtttaagcagccagcagtcggtgttcttggtgattctctccagaatactctc 60
|||||
DB 450 ATGGATTAAAGACCGCACACAGTCGGTGTTCGTAATTCCTCCAGAAATCACTGCT 391
|||||
QY 61 tcaagataatggtccagcagaacttcagccagcagccttcaactcaagcccttgcaa 120
|||||
DB 390 TCAGAAATATGATCCACAGAACTTTACAGCACTTTTCAACACAAAGCCCTTGCA 331
|||||
QY 121 aaattattctgtagaanaaacttaagagactatccagatctggttggaattatg 180
|||||
DB 330 AAATATATTTGCTAGAAAATCTTAAGGGATATCCATTCTGGGGCTCTGTTTGT 271
|||||
QY 181 acccttctcttgagagatctctcttcaactgttaaacatatacgaagttctcc 240
|||||
DB 270 TCATTAAATTTGAGACCTCTCTTAATTGCAATGAAAAGAAAACCAACAACCTGTGATCA 211
|||||
QY 241 ttataattctctcagagatccatctcgtggtgttcttcatatattctcgagcc 300
|||||

```

```

DB 210 AAACATCAATTTGTGTTATCTCTACCAAAATAGTCAGTGAAGCGTGTACTGCTGTT 151
|||||
QY 301 ttccctaatgtagaagaanaaaccaagaactctgggaatttattacattagtg 360
|||||
DB 150 CT-----TGGGAATTTTGATTAATCAATGATG 126
|||||
QY 361 acttcagcatatgaattatcattctctgccttcttcaatttgggtgacctca 420
|||||
DB 125 ACTTCAGCATTAATGAAATTAATTCATTCTCTGCTTCTCAATTTGGGGTCCACATCA 66
|||||
QY 421 gagagattgtagtggaacaatggttgga 450
|||||
DB 65 GAGGATGTGATTTGCAACAATGTTGTTGA 36
|||||
RESULT 6
AA707529 281 bp mRNA EST 12-JAN-1999
LOCUS a141a12.s1 Soares_testis_NHT Homo sapiens cDNA clone 1292062 3',
DEFINITION mRNA sequence.
ACCESSION AA707529
VERSION AA707529.1 GI:2717447
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 281)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TUMOR National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
FEATURES
SOURCE
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
Insert Length: 456 Std Error: 0.00
Seq primer: -40m13 fwd. EF from Amersham
High quality sequence stop: 218.
Location/Qualifiers
1..281
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1292062"
/clone_id="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTTCACATCTGAGTGAAGGAGCGCGCCCAATTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 94 a 54 c 45 g 88 t
ORIGIN
Query Match 43.8%; Score 197.2; DB 10; Length 281;
Best Local Similarity 98.5%; Pred. No. 2.1e-34;
Matches 199; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 249 tcttcagagatccatctcgtggtgttcttcaattatctgagccttccaat 308
|||||

```

Db 237 TCTTAGGGATATCCATCTCGGGCTCTGTTTGTTCATTAAATTCGAGCCCTTCTAA 178
 QY 309 tgcagtgaaagaaagaaacacagaaactctggaatttgattgaacttgacttcag 368
 Db 177 TGAGAGAAAGAAAGAAACACAGAAACTCTGGAAATTTGATTACATGATGACTTTGAG 118
 QY 369 cactatgataattcattctctccttccttcactatcttgagggtgcgaactcaagagattg 428
 Db 117 CATATATTGAATTAATTCATTTCTCTGCTTTTCATTTTGGGTGCACACAGAGATTG 58
 QY 429 tgatgtgacaactgttgtga 450
 Db 57 TGATGTGAACATGTTGTGA 36

RESULT 7
 AA470059 415 bp mRNA EST 09-NOV-1997
 LOCUS 2194h05.r1 Soares testis NHT Homo sapiens cDNA clone IMAGE:730041
 DEFINITION 5', similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ; ,
 mRNA sequence.

ACCESSION AA470059.1 GI:2197368
 VERSION AA470059.1
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 415)
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisler,G., Jost,S.,
 Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,
 J., Moore,B., Schellenberg,K., Stepien,M., Tan,F., Theisling,B.,
 White,Y., Wylie,T., Waterston,R. and Wilson,R.
 JOURNAL Unpublished (1997)
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.jnl.gov) for further information.
 Insert Length: 656 Std Error: 0.00
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 393.

FEATURES
 source
 Location/Qualifiers
 1..415

/organism="Homo sapiens"
 /db_xref="GDB:5926570"
 /db_xref="taxon:9606"
 /clone="IMAGE:730041"
 /clone_1lb="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pRT73D-Pac (Pharmacia) with a modified
 polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech laboratories
 , Inc., and primed with a Not I - oligo(dT) primer [5'
 TCTTACCATCTGAGAGTGGAGCGGCCGCCCAATTTTCTTTTCTTTT 3']
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pRT73 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bernaldo."

BASE COUNT 129 a 96 c 67 g 123 t

Query Match 35 1%; Score 157.8; DB 10; Length 415;
 Best Local Similarity 98.8%; Pred. No. 1.3e-25;
 Matches 159; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 atgagattcaagcaccgcacagctccggtgttctgtattcttcacgaatcactgct 60
 Db 57 ATGAGATTCAAGCACCCGACACACTCCGGTGTTCCTGTATTTCCTCCAGAAATCACTGCT 116
 QY 61 tcagaataigagtcacagaaacttcagccagacattcaactcaagcccccttgcaa 120
 Db 117 TCAGAAATGAGTCCACAGAACTTTGAGCCAGACCTTTTCAACTCAAGCCCTTGCAA 176
 QY 121 aaattattgctaganaaatgaaatccttaggacatcca 161
 Db 177 AAATTAATTTGCTAGAAAAATGAAAAATCTTAGGGGATATCA 217

RESULT 8
 BE107659/c 484 bp mRNA EST 13-JUN-2000
 LOCUS UI-R-BT1-ame-d-04-0-UI.s1 UI-R-BT1 Rattus norvegicus cDNA clone
 DEFINITION UI-R-BT1-ame-d-04-0-UI 3', mRNA sequence.
 ACCESSION BE107659
 VERSION BE107659.1 GI:8499769
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 484)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.iowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to identify it as a clone from the
 normalized testis library cDNA Library Preparation: M.B. Soares Lab
 Clone distribution: clones will be available through Research
 Genetics (www.resgen.com)
 Seq primer: M13 Forward
 POLA=Yes.

FEATURES
 source
 Location/Qualifiers
 1..484

/organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-BT1-ame-d-04-0-UI"
 /clone_1lb="UI-R-BT1"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pRT73D-Pac (Pharmacia) with a modified
 polylinker. Site.1: Not I; Site.2: Eco RI; The library
 UI-R-BT1 is a subtracted library derived from a mixture of
 the following tissues: hippocampus, thalamus, mid-brain,
 medulla, corpus striatum, cerebral cortex and testis. For
 a detailed description of the library from which this
 clone was derived, please visit our web site at
 ratest.eng.uiowa.edu. The subtraction has been previously
 described in (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)
 TAG_LIB=UI-R-BT1
 TAG_TISSUE=testis
 TAG_SEQ=ACGCAG"

BASE COUNT 165 a 98 c 96 g 125 t

Query Match 32.4%; Score 145.8; DB 10; Length 484;
 Best Local Similarity 82.0%; Pred. No. 6.3e-23;
 Matches 168; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

DB 161 agatccgttggaattacacattcttcttgagggtatctctcttccattccattgtaaa 220
 |||||
 DB 484 AGATCCGTGGAATTACACATTCTTCTGAGGATGTTGTCCTTTCCATCCCTGGTAA 425
 |||||

DB 221 aacacatcaagagttcccttatacttcttcagagatccattccattcgggctctgtt 280
 |||||
 DB 424 ACCATFACCAAGTTCCTTTATATTTCTCTCAGAGATCTTTCTGCGGCTCTGCTT 365
 |||||

DB 281 tgttataatctcggagccttcccaattgcagtgagaaagaaacacagaactctgg 340
 |||||
 DB 364 TGTTCATTAACCTCTGGACCTTTCTGATTTGCTGAAAGAAACACAGACTCTGA 305
 |||||

DB 341 gaatttgattacattgattgtt 365
 |||||
 DB 304 TAAAAATGAGCCAGCGGATGATTT 280
 |||||

RESULT 9
 BE638325/c 224 bp mRNA EST 28-AUG-2000
 LOCUS EST00022 mouse testis apoptosis subtraction cDNA library Mus
 DEFINITION BE638325
 ACCESSION BE638325
 VERSION BE638325.1 GI:9932068
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 224)
 JIANG, H., LI, L. and LU, G.
 TITLE Expressed sequence tags from mouse testis apoptosis subtraction
 cDNA library
 JOURNAL Unpublished (2000)
 COMMENT Contact: Jiang H
 Department of Pharmacology
 Hunan Medical University
 88 Xiangya road, Changsha, Hunan 410078, China
 Tel: 086-0731-4405036
 Fax: 086-0731-4497661
 Email: 1j12epublic.cs.hn.cn
 Seq primer: T7 Promoter Primer
 High quality sequence stop: 224.

FEATURES
 source
 1..224
 /organism="Mus musculus"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="MTA22"
 /clone_lib="mouse testis apoptosis subtraction cDNA
 library"
 /sex="male"
 /tissue_type="testis"
 /cell_type="spermatogenesis cells"
 /dev_stage="adult"
 /lab_host="JM109"
 /note="Vector: pUCm-T"
 BASE COUNT 58 a 40 c 61 g 65 t
 ORIGIN

Query Match 28.0%; Score 126.2; DB 10; Length 224;
 Best Local Similarity 76.4%; Pred. No. 1.8e-18;
 Matches 155; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

DB 25 ccgagtttcggtattccctcagaatcgtcttcagatgatgagccagaactt 84
 |||||

DB 204 CCTGTTCCTGGTGTTCCTTCAGAGGTCACCACTCAAGATACCAACAGACTT 145
 |||||
 DB 85 tcagccagcacttcccaactcaagaagcccttgcaaaaattatctgtagaanaatgaa 144
 |||||

DB 144 ACAGGCACAGCCTATTAATCCCAAAACGAGTCAGAGATCATTTATTAAGAAATGGAG 85
 |||||

DB 145 atcttaggactatccagatcctgtttggaattatgaccttcttggagttatctt 204
 |||||

DB 84 ACTTAGGACACACCAACCAATCTGTTGGAAATCAATGAACTTTCAATTTGAGATTTTC 25
 |||||

DB 205 ctttcacctgttaaaccata 227
 |||||

DB 24 CTTTTCACCTTGTTAAACCATTA 2
 |||||

RESULT 10
 AA435988/c 382 bp mRNA EST 09-NOV-1997
 LOCUS AA435988
 DEFINITION z003a08.s1 Soares_testis_NHT Homo sapiens cDNA IMAGE:730742
 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCTYE ANTIGEN CD20 ;,
 mRNA sequence.
 ACCESSION AA435988
 VERSION AA435988
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 382)
 HILLIER, L., ALLEN, M., BOWLES, L., DUBUQUE, T., GEISEL, G., JOST, S.,
 KRIZMAN, D., KUCADA, F., LACY, M., LE, N., LEMON, G., MARTA, M., MARTIN
 WHITE, Y., WYLLIE, T., WATERSTON, R. and WILSON, R.
 TITLE WashU-NCI human EST Project
 JOURNAL Unpublished (1997)
 COMMENT Contact: Wilson Rk
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@imgl.gov) for further information.
 Possible reversed clone: similarity on wrong strand
 Insert Length: 733 Sca Error: 0.00
 Seq primer: -41m13 fwd. EF from Amersham
 High quality sequence stop: 280.

FEATURES
 source
 1..382
 /organism="Homo sapiens"
 /db_xref="GDB:592743"
 /db_xref="taxon:9606"
 /clone="IMAGE:730742"
 /clone_lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pRT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc., and primed with a Not I - oligo(dT) primer [5',
 TGTTCACATCTCAAGAGGAGGCGGCCCAATTTTATTTTATTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pRT73 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bernaldo."

BASE COUNT 135 a 73 c 67 g 107 t
 ORIGIN

Query Match 25.1%; Score 113; DB 10; Length 382;
 Best Local Similarity 100.0%; Pred. No. 1.5e-15;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 338 tgggaatttgatattacattgacttcacagattatgaattatattccttcctcctt 397
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 150 TGGGAATTTTGATTTACATGATGACTTTCAGCATTTATTAATTAATTCATTTCTCTGCTT 91
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 398 ttccaatttgggtgcacacagagattgtgattgtgaacaatgtgttga 450
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 90 TCTCAATTTTGGGGTCCACCTCAGAGATTGTGATTTGTCACATCTGTGTA 38
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 11
 AA416972 387 bp mRNA EST 09-NOV-1997
 LOCUS AA416972/c zt94h05.s1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:730041
 DEFINITION 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;,
 mRNA sequence.
 ACCESSION AA416972.1 GI:2077080
 VERSION AA416972
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 387)
 Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
 White,X., Wylie,T., Waterscon,R. and Wilson,R.
 WashU-NCI human EST Project
 UNPUBLISHED (1997)
 TITLE JOURNAL
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 656 Std Error: 0.00
 Seq primer: 41m13 fwd. ET from Amersham.
 FEATURES
 source Location/Qualifiers
 1..387
 /organism="Homo sapiens"
 /db_xref="GDB:5926570"
 /db_xref="taxon:9606"
 /clone="IMAGE:730041"
 /clone_11b="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc., and primed with a Not I - oligo(dT) primer [5'
 TGTTCACATCTGAGTGGAGCGGCCCAATTTTATTTTATTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to Cols5, and was
 constructed by Bento Soares and M. Fatima Bonaudo."
 BASE COUNT 138 a 77 c 67 g 105 t
 ORIGIN
 Query Match 25.1%; Score 113; DB 10; Length 387;
 Best Local Similarity 100.0%; Pred. No. 1.5e-15;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 338 tgggaatttgatattacattgacttcacagattatgaattatattccttcctcctt 397
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 145 TGGGAATTTTGATTTACATGATGACTTTCAGCATTTATTAATTAATTCATTTCTCTGCTT 86
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 398 ttccaatttgggtgcacacagagattgtgattgtgaacaatgtgttga 450
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 85 TCTCAATTTTGGGGTCCACCTCAGAGATTGTGATTTGTCACATCTGTGTA 33
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 12
 A0108532 538 bp DNA GSS 29-AUG-1998
 LOCUS A0108532 CIT-HSP-237967.TF CIT-HSP Homo sapiens genomic clone 237967, DNA
 DEFINITION sequence.
 ACCESSION A0108532.1 GI:3485222
 VERSION A0108532
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 538)
 Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
 Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
 Venter,J.C.
 Use of a random human BAC End Sequence Database for Sequence-Ready
 Map Building
 UNPUBLISHED (1998)
 COMMENT Other_GSSs: CIT-HSP-237967.TF
 CONTACT Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13-21
 Class: BAC ends.
 FEATURES
 source Location/Qualifiers
 1..538
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="237967"
 /clone_11b="CIT-HSP"
 /sex="Male"
 /cell_type="Sperm"
 /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
 HindIII"
 BASE COUNT 140 a 108 c 115 g 175 t
 ORIGIN
 Query Match 25.0%; Score 112.4; DB 13; Length 538;
 Best Local Similarity 99.1%; Pred. No. 1.9e-15;
 Matches 113; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 337 ctgggaatttgatattacattgacttcacagattatgaattatattccttcctcctt 396
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 168 CAGGGAATTTTGATTTACATGATGACTTTCAGCATTTATTAATTAATTCATTTCTCTGCTT 227
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 397 ttccaatttgggtgcacacagagattgtgattgtgaacaatgtgttga 450
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 228 TTCTCAATTTTGGGGTCCACCTCAGAGATTGTGATTTGTCACATCTGTGTA 281
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 13
 AA758635 394 bp mRNA EST 29-DEC-1998
 LOCUS AA758635/ah67b04.s1 Soares testis_NHT Homo sapiens cDNA clone 1320655 3'
 DEFINITION similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;,
 mRNA sequence.
 ACCESSION AA758635.1 GI:2806498
 VERSION AA758635
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 394)
 NC1-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NC1-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/BLND at:
www.bio.livnl.gov/dbp/image/image.html
 Insert length: 661 Std Error: 0.00
 Seq primer: -40m13 fwd. Er from Amersham
 High quality sequence stop: 390.
 Location/Qualifiers

FEATURES

SOURCE

1. 394
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="1320655"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: p7T3D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCCGCCCAATTTTCTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

133 a 73 c 67 g 121 t

ORIGIN

Query Match 21.0%; Score 94.6; DB 10; Length 394;
 Best Local Similarity 88.0%; Pred. No. 2e-11;
 Matches 103; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 249 tcttcagagataccatctcgtggtcgttctgtcattatcttgagccttcctaat 308
 394 TCTTTCAGGATATCCATTCCTGGGCTCTGTATGTATTAATTCCTGAAGCTTCTTAAT 335
 309 tgcagtgagaagaacacacagaaactctggaatttgattacatgatgactt 365
 334 TGCAGTGAAGAAAGAAACACAGAAACTCTGATTAATTAATGAGCCGATTAATGAAATT 278

RESULT 14
 BE638321 218 bp mRNA EST 28-AUG-2000
 LOCUS BE638321 mouse testis apoptosis subtraction cDNA library Mus
 DEFINITION Musculus cDNA clone M7A13 5', mRNA sequence.
 ACCESSION BE638321
 VERSION BE638321.1 GI:9932064
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 218)
 AUTHORS Jiang, H., Li, L. and Lu, G.
 TITLE Expressed sequence tags from mouse testis apoptosis subtraction cDNA library
 JOURNAL Unpublished (2000)
 COMMENT Contact: Jiang H
 Department of Pharmacology

Human Medical University
 88 Xiangya road, Changsha, Hunan 410078, China
 Tel: 086-0731-4805036
 Fax: 086-0731-4497661
 Email: lj12@public.cs.hn.cn
 Seq primer: T7 Promoter Primer
 High quality sequence stop: 218.
 Location/Qualifiers

FEATURES

SOURCE

1. 218
 /organism="Mus musculus"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="M7A13"
 /sex="male"
 /lab_host="mouse testis apoptosis subtraction cDNA library"
 /tissue="testis"
 /cell_type="spermatogenesis cells"
 /dev_stage="adult"
 /lab_host="JMI09"
 /note="Vector: pUCm-T"

BASE COUNT

61 a 39 c 46 g 72 t

ORIGIN

Query Match 18.4%; Score 83; DB 10; Length 218;
 Best Local Similarity 79.7%; Pred. No. 8.7e-09;
 Matches 98; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 104 ctcaagcccttggaanaaatatttgcctagaagaataaacttaggactatccaga 163
 DB 198 CCCAAACCGAGTGCAGAGATCATTAATTAAGAAATTTGGAGATCTTAGGACACCCAGA 139
 QY 164 tctcttggaattatgaccttctcttcttggaagtatctcttcacactgttaaac 223
 DB 138 TCTCTTTGGAATCATGAACATTTCATTTTGAGAGTCATTTCTTTCACCTTGGTAAC 79
 QY 224 cat 226
 DB 78 CCT 76

RESULT 15
 BE6842 421 bp DNA GSS 09-APR-1999
 LOCUS RPI11-26F10.TP RPI1-11 Homo sapiens genomic clone RPI1-11-26F10,
 DEFINITION DNA sequence.
 ACCESSION BE6842
 VERSION BE6842.1 GI:2927974
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 421)
 AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Sun, E., Wible, C., de Jong, P. and Venter, J.C.
 TITLE Use of BAC end sequences for sequence-ready Map Building (1998)
 JOURNAL Unpublished (1998)
 COMMENT Other GSSs: RPI11-26F10.TV
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are derived from the human BAC library RPI1-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genetics (info@resgen.com). BAC end search page:

http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html

Seq primer: SP6
Class: BAC ends.

FEATURES
source location/Qualifiers

1. .421
/organism="Homo sapiens"
/db_xref="GDB:7509729"
/db_xref="taxon:9606"
/clone="RPC1-11-26F10"
/clone_lib="RPC1-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC11 Human Male BAC Library"
BASE COUNT 124 a 95 c 71 g 131 t
ORIGIN

Query Match 15.7%; Score 70.6; DB 13; Length 421;
Best Local Similarity 94.8%; Pred. No. 4.6e-06;
Matches 73; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 148 ttaggactatccagatccctgttggaaatagaccttcttcttggagttatctcctt 207
|| |||||
DB 345 TTGCAGACTATCCAGATCCTGTTGGAAATATGACCTTTCTTTGGAGTATCTCCTT 404
QY 208 ttaccttggttaaac 224
|||||
DB 405 TTCACCTTGTGTTAAAC 421

Search completed: November 29, 2001, 00:18:42
Job time: 9879 sec

THIS PAGE BLANK (USPTO)